



Tıp Fakültesi

T.C.

MARMARA ÜNİVERSİTESİ TIP FAKÜLTESİ

ÇOCUK SAĞLIĞI VE HASTALIKLARI ANABİLİM DALI

**ERKEN BAŞLANGIÇLI BAĞIRSAK
HASTALIKLARINDA PROTEİN KAYBETTİRİCİ
ENTEROPATİ VE İNFLAMATUAR BAĞIRSAK
HASTALIKLARININ AYIRT EDİCİ VE ORTAK
ÖZELLİKLERİ**

DR. SELMAN BİRKAN ÖZER

UZMANLIK TEZİ

İSTANBUL 2023



Tıp Fakültesi

T.C.

MARMARA ÜNİVERSİTESİ TIP FAKÜLTESİ

ÇOCUK SAĞLIĞI VE HASTALIKLARI ANABİLİM DALI

**ERKEN BAŞLANGIÇLI BAĞIRSAK
HASTALIKLARINDA PROTEİN KAYBETTİRİCİ
ENTEROPATİ VE İNFLAMATUAR BAĞIRSAK
HASTALIKLARININ AYIRT EDİCİ VE ORTAK
ÖZELLİKLERİ**

DR. SELMAN BİRKAN ÖZER

UZMANLIK TEZİ

DANIŞMAN: PROF. DR. AHMET OĞUZHAN ÖZEN

İSTANBUL 2023

ÖNSÖZ

Tüm eğitim hayatım boyunca bana yol gösteren ve emek harcayan değerli öğretmenlerime;

Tez çalışmam için yardımlarını esirgemeyen, sorularımı her daim cevaplayan, bilgi ve tecrübeleri ile yolumu aydınlatan, başta tez danışmanım Prof. Dr. Ahmet Oğuzhan ÖZEN olmak üzere, Marmara Üniversitesi Tıp Fakültesi Çocuk Allerji ve İmmünoloji Bilim Dalı'nın tüm üyelerine;

İstatistiksel analizler ile makine öğrenimi yöntemlerinin seçimi, kontrolü ve yorumlanması konularında, değerli vaktini ayırarak bana yardımcı olan Prof. Dr. Esra AKDENİZ'e;

Uzmanlık eğitimim süresince hep yanımda olan eğitim danışmanım Prof. Dr. Hülya Selva BİLGİN'e; benimle beraber bu yolu yürüten tüm eşkıdemlerime; gece gündüz beraber çalıştığım tüm mesai arkadaşlarıma;

Varlıkları ile bana güç veren sevgili eşim Merve ÇAĞLAR ÖZER'e, oğlum Atahan ÖZER'e, anneme, babama, kardeşime ve en az onlar kadar değerli olan eşimin annesi, babası ve kardeşine;

Teşekkür ederim.

ÖZET

Giriş: İnflamatuvar Bağırsak Hastalıkları (İBH), genellikle multifaktöriyel bir hastalık olmakla birlikte çocuklarda, özellikle 6 yaşından küçük olanlarda ortaya çıkan bazı formları monogenik etyolojiye sahiptir (mİBH) ve 100’den fazla çeşit genle ilişkilendirilmiştir. İmmün sistemin bir ya da daha fazla elemanını tutan bozukluklar (Doğuştan Bağışıklık Kusurları, DBK) ve / veya barsakların yapısal elemanlarına ait anormallikler mİBH ile başvuruya neden olabilir. Altta yatan mekanizmaların belirlenmesi tedaviyi şekillendirme bakımından yüksek öneme sahiptir. mİBH’nın moleküler etyolojileri son 10 yılda yeni yeni aydınlatılmaya başlamış olup farklı mekanizmalarla ortaya çıkan hastalıkların (etkilenmiş hücre ve moleküler yolaklar) fenotipik spektrumları henüz tam olarak belirlenmemiştir. Ağırlıklı olarak gastroenteroloji ve klinik immünoloji hekimlerince izlenen mİBH’lar pekçok sistemi etkileyen anormallikler sergilediğinden, tanı ve tedavi süreçleri multidisipliner ve bütüncül bakış açısı ile ele alınmalıdır.

Amaç: Bu çalışmada moleküler etyolojileri aydınlatılmış ya da kuşkulu mİBH olgularının klinik, immünolojik ve genetik özellikleri ayrıntılı olarak değerlendirilmiş ve insan fenotip ontolojisi kullanılarak semptomatoloji ve patolojik bulguların spektrumu gözden geçirilmiştir. Güncel mİBH sınıflandırma sistemleri ve DBK sınıflandırmaları kullanılarak farklı kategorilerdeki hastaların fenotipik benzerlikleri ve ayrıştıkları noktalar incelenmiştir. Makine öğrenmesi algoritmaları ile klinik bulgular ve immünolojik testlerin moleküler tanı ve hastalık kategorilerini öngörme potansiyelleri araştırılmıştır.

Yöntem: Marmara Üniversitesi Çocuk Alerji-İmmünoloji Bilim Dalı’nın öncülüğünde yürütülen ulusal mİBH çalışmasına dahil edilmiş olguların dosyaları gözden geçirilmiştir. Çalışma evreni farklı kohortlardan oluşmaktadır: Primer protein kaybettirici enteropati (PKE) tanılı olgular, immün disregülasyon kohortunda çalışılan ve gastrointestinal bulguları olan hastalar, immün yetmezlik ile takip edilen ve eş zamanlı gastrointestinal bulguları olan hastalar. Rastgele Orman ve Gradyan Arttırma algoritmaları kullanılarak farklı parametrelerin mİBH ve DBK

sınıflandırmalarını öngörme başarısı test edilmiştir.

Bulgular: Çalışma için 520 dosya incelenmiş ve bunlardan 200'ü ileri inceleme için seçilmiştir. Seçilen hastaların %36'sının kız, %64'ünün erkek olduğu görülmüştür. Hastaların %61'inde akraba evliliği tespit edilmiştir. Hastaların %66'sında genetik mutasyon saptanmıştır. DBK sınıflamasının neredeyse tüm sınıflarından, toplamda 53 farklı genetik mutasyon görülmüştür. Genetik mutasyonu tespit edilen hastaların %80'inde mİBH ilişkili genetik mutasyon gözlenmiştir. Hastaların immün fenotiplerinde, ait oldukları sınıfları işaret eden motifler olduğu görülmüştür. Bu motiflerin aynı sınıfın ishal ile seyretmeyen mutasyonları ile benzer olduğu görülmüştür. Çalışmadaki hastalarda, geniş sayıda organ ve sistemde, toplamda 481 farklı fenotipik özellik gözlenmiştir. Hasta başına tespit edilen fenotipik özellik ortancasının 19 (IQR 14 - 24)'dur. Fenotipik özellikler kullanılarak yapılan makine öğrenim denemelerinde, hastanın tanısına işaret edebilecek özellikler tespit edilmiştir.

Sonuç: Klinik öykü, fizik inceleme, temel biyokimyasal testler, fenotipe özgü planlanan akım sitometri çalışmaları hastalık kategorilerini kabaca belirleyebilse de klinik spektrumun genişliği ve sınıflar arası örtüşmeler moleküler tanıyı ön görecektir. hassas bir modele olanak tanımamaktadır. Öte yandan saptadığımız bazı fenotipik ontoloji terimlerindeki kümelenmeler dikkat çekici bulunmuştur. Egzom ve genom sekanslamaları gibi genetik incelemeler erken dönemde planlanmalıdır.

Anahtar kelimeler: monogenik İnflamatuvar Bağırsak Hastalıkları, Primer İmmün Yetmezlik, Protein Kaybettirici Enteropati, HPO, yapay zeka, WGS, WES

ABSTRACT

Introduction: Inflammatory Bowel Diseases (IBD), although typically multifactorial in nature, can manifest as specific monogenic forms (mIBD) in children, particularly those under the age of 6. These monogenic forms are linked to over 100 different genes.. Conditions impacting the immune system's components, known as Primary Immunodeficiencies (PIDs), as well as abnormalities in the structural elements of the intestines, can contribute to the development of mIBD. Understanding these underlying mechanisms is of paramount importance in guiding treatment strategies. The molecular etiologies of mIBD have only begun to be elucidated in the last decade, and the phenotypic spectrums of diseases arising from different mechanisms (affected cells and molecular pathways) have not yet been fully determined. mIBDs are primarily managed by gastroenterologists and clinical immunologists, but it is crucial to recognize that these conditions involve abnormalities affecting multiple systems. Therefore, a multidisciplinary and holistic approach should be employed in both the diagnosis and treatment processes.

Objective: In this study, we conducted a comprehensive evaluation of the clinical, immunological, and genetic features in cases with confirmed or suspected molecular etiologies of mIBD. We systematically reviewed the spectrum of symptoms and pathological findings using human phenotype ontology, analyzing both the commonalities and distinctions among patients falling into various categories according to the current mIBD and PID classifications. Furthermore, we explored the feasibility of machine learning algorithms to predict molecular diagnoses and disease categories by leveraging clinical data and immunological test results.

Method: We conducted a review of the medical records of individuals who were part of the national mIBD study conducted by the Department of Pediatric Allergy and Immunology at Marmara University. The study population is diverse and includes several cohorts: individuals diagnosed with Primary Protein-Losing Enteropathy (PKE), patients from the immune dysregulation cohort presenting with gastrointestinal symptoms, and patients with concurrent immunodeficiency and

gastrointestinal symptoms who are under medical follow-up. Using Random Forest and Gradient Boosting algorithms, the predictive success of different parameters for mIBD and PID classifications has been tested.

Results: While the clinical history, physical examination, basic biochemical tests, and phenotype-specific targeted flow cytometry studies can roughly determine disease categories, the breadth of the clinical spectrum and the overlaps between classes do not allow for a precise model to predict molecular diagnosis. On the other hand, we have found certain clusters in the phenotypic ontology terms we identified to be noteworthy. Genetic investigations such as exome and genome sequencing should be planned early on.

Keywords: Monogenic Inflammatory Bowel Diseases, Primary Immunodeficiency, Protein-Losing Enteropathy, HPO (Human Phenotype Ontology), artificial intelligence, Whole Genome Sequencing (WGS), Whole Exome Sequencing (WES)

İÇİNDEKİLER

ÖNSÖZ	iv
ÖZET	v
ABSTRACT	vii
İÇİNDEKİLER	ix
KISALTMALAR	x
TABLolar VE ŞEKİLLER	xii
1. GİRİŞ VE AMAÇ	1
2. GENEL BİLGİLER	5
2.1. Protein Kaybettirici Enteropati	5
2.2. Primer İntestinal Lenfanjektazi	6
2.3. İnflamatuvar Bağırsak Hastalıkları	6
2.4. Erken ve Çok Erken Başlangıçlı İnflamatuvar Bağırsak Hastalıkları	8
2.5. Monogenik İnflamatuvar Bağırsak Hastalıkları	9
2.6. İmmün Yetmezlikler	12
2.7. İmmün Yetmezlikler ve Monogenik İnflamatuvar Bağırsak Hastalıkları	16
2.8. İnsan Fenotip Ontolojisi (Human Phenotype Ontology, HPO)	19
2.9. Yapay Zeka ve Makine Öğrenimi	23
2.9.1. Karar Ağaçları (Decision Trees - DT) Algoritması	25
2.9.2. Rastgele Orman (Random Forest - RF) Algoritması	27
2.9.3. Gradyan Artırmalı Karar Ağaçları (Gradient Boosting Decision Trees - GBDT) Algoritması	29
2.9.4. Makine Öğrenimi Modellerini Değerlendirme Ölçütleri	30
2.9.5. Veri Dağılımı ve Eksik Veriler	33
2.10. Genetik	35
3. GEREÇ VE YÖNTEM	40
3.1. Araştırma Evreni	40
3.2. İzinler ve Onamlar	41
3.3. Veri Toplama	41
3.4. Analiz	42
4. BULGULAR	44
5. TARTIŞMA	86
6. KAYNAKÇA	90
7. EKLER	102

KISALTMALAR

- AUC: *Area Under Curve* (Eğri Altında Kalan Alan)
- CD: *Cluster of Differentiation* (Başkalaşım Kümesi)
- CH: Crohn Hastalığı
- ÇEBİBH: Çok Erken Başlangıçlı İnflamatuvar Bağırsak Hastalığı
- DBK: Doğuştan Bağışıklık Kusurları
- DNA: Deoksiribonükleik Asit
- DT: *Decision Trees* (Karar Ağaçları)
- EBİBH: Erken Başlangıçlı İnflamatuvar Bağırsak Hastalığı
- FISH: *Fluorescence in situ Hybridization* (Yerinde Floresan Melezleme)
- FN: *False Negative* (Yanlış Negatif)
- FP: *False Positive* (Yanlış Pozitif)
- GBDT: *Gradient Boosting Decision Trees* (Gradyan Artırmalı Karar Ağaçları)
- GİS: Gastrointestinal Sistem
- GO: *Gene Ontology*
- GWAS: *Genome-wide Association Studies* (Genom Boyu İlişki Çalışmaları)
- HLH: Hemofagositik lenfhistiyositoz
- HPO: *Human Phenotype Ontology* (İnsan Fenotip Ontolojisi)
- HSP: Henoch-Schönlein Purpurası
- HÜS: Hemolitik-Üremik Sendrom
- IEI: *Inborn Errors of Immunity* (Bağışıklığın Doğuştan Bozukluğu)
- Ig: İmmunoglobulin
- IUIS: *International Union of Immunological Societies* (Uluslararası İmmünoloji Dernekleri Birliği)
- İBH: İnflamatuvar Bağırsak Hastalıkları
- İBİBH: İnfantil Başlangıçlı İnflamatuvar Bağırsak Hastalığı
- İL: İntestinal Lenfanjektazi
- IVIG: *Intravenous Immunoglobulin* (İntravenöz İmmunoglobulin)
- IQR: *Interquartile Range* (Çeyrekler Arası Aralık)

K-S: Kolmogorov-Smirnov testi
K-W: Kruskal-Wallis testi
MAS: Makrofaj aktivasyon sendromu
MCT: Orta Zincirli Yağ
NGS: *Next-Generation Sequencing* (Yeni Nesil Dizileme)
NK: *Natural Killer* (Doğal Öldürücü)
OBO: *The Open Biological and Biomedical Ontologies Foundry*, daha önceden *Open Biomedical Ontologies Foundry* (Açık Biyolojik ve Biyomedikal Ontolojiler Atölyesi, daha önceden Açık Biyomedikal Ontolojiler Atölyesi)
OMIM: Online Mendelian Inheritance in Man Database (İnsan Mendeliyen Kalıtları Çevrimiçi Veritabanı)
PCR: *Polymerase Chain Reaction* (Polimeraz Zincir Reaksiyonu)
PİL: Primer İntestinal Lenfanjektazi
PİY: Primer İmmün Yetmezlik
PKE: Protein Kaybettirici Enteropati
RCPCH: *The Royal College of Paediatrics and Child Health* (Pediatri ve Çocuk Sağlığı Kraliyet Koleji)
RF: *Random Forest* (Rastgele Orman)
ROC : *Receiver Operating Characteristic* (Alıcı İşletim Karakteristiği)
RTE: *Recent Thymic Emigrant* (Yeni Timik Göçmen)
S-W: Shapiro-Wilk testi
SS: Standart sapma
sİBH: Sınıflandırılmayan İBH
SLE: Sistemik Lupus Eritematozus
SMOTE: *Synthetic Minority Over-sampling Technique* (Azınlığın Yapay Aşırı Örnekleme Tekniği)
SNP: *Single Nucleotide Polymorphism* (Tek Nükleotid Polimorfizmi)
TCR: *T Cell Receptor* (T Hücre Reseptörü)
TN: *True Negative* (Gerçek Negatif)
TNR: *True Negative Rate* (Gerçek Negatif Oranı)
TP: *True Positive* (Gerçek Pozitif)
TPR: *True Positive Rate* (Gerçek Pozitif Oranı)

USG: Ultrasonografi

ÜK: Ülseratif Kolit

WES: *Whole Exome Sequencing* (Tüm Ekzom Dizileme)

WGS: *Whole Genome Sequencing* (Tüm Genom Dizileme)

TABLolar VE ŐEKİLLER

Tablo 1: DBK Sınıflamasının özetİ (Mayıs 2022)	14
Tablo 2: PİY için uyarıcı bulgular	15
Tablo 3: Tüm kohortlarda genetik etyoloji dağılımı	44
Tablo 4: Genlerin DBK ve mİBH sınıflamalarına göre dağılımı	45
Tablo 5: Hastaların demografik dağılımları	46
Tablo 6: Araştırma evrenindeki hastalarda görölen genetik mutasyonlar ve sıklıkları	47
Tablo 7: Araştırma evrenindeki hastaların genetik mutasyonlarının farklı sınıflamalara göre dağılımı	49
Tablo 8: DBK sınıflamasına göre ayrılan gruplarda boy z skoru, kilo z skoru ve semptom başlangıç yaşı dağılımı	52
Tablo 9: PKE gruplamasına göre ayrılan gruplarda boy z skoru, kilo z skoru ve semptom başlangıç yaşı dağılımı	52
Tablo 10: Araştırma evrenindeki tüm HPO terimleri ve görölme sıklıkları	67
Tablo 11: Makine öğreniminde kullanılan deęişkenler	78
Tablo 12: Makine öğrenimi modellerinin özellikleri ve sonuçları	79
Tablo 13: Birleştİrilmiř HPO terimlerinin farklı DBK sınıflarındaki karşılıkları	81
Őekil 1: mİBH düşünölen olgulara yaklaşım.	11
Őekil 2: PİY řüpheli İBH olgularına yaklaşım.	18
Őekil 3: HPO ağacından bir kesit	20
Őekil 4: Karar ağacı örneęi	26
Őekil 5: Rastgele orman algoritması örneęi	28

Şekil 6: Araştırmaya dahil edilen hastaların seçildiği kaynaklar	40
Şekil 7: İnteraktif grafikler için QR kodu	43
Şekil 8: Araştırma evrenindeki hastaların genetik etyolojilerinin farklı sınıflamalara göre dağılımı	48
Şekil 9: Semptom başlama yaşı, boy z skoru ve kilo z skoru dağılımları	50
Şekil 10: DBK sınıflamasına göre ayrılan gruplarda boy ve kilo z skoru dağılımı	50
Şekil 11: DBK sınıflamasına göre ayrılan gruplarda semptom başlama yaşı dağılımı	51
Şekil 12: PKE gruplamasına göre ayrılan gruplarda boy ve kilo z skoru dağılımı	53
Şekil 13: DBK sınıflamasına göre ayrılan gruplarda semptom başlama yaşı dağılımı	53
Şekil 14: Ig seviyelerinin yaşa göre durumunun DBK sınıflamasına göre dağılımı	54
Şekil 15: Hemogramın yaşa göre durumunun DBK sınıflamasına göre dağılımı	55
Şekil 16: Basit immünofenotiplemede hücre oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı	56
Şekil 17: Basit immünofenotiplemede hücre sayılarının yaşa göre durumunun DBK sınıflamasına göre dağılımı	57
Şekil 18: B Lenfosit alt gruplarının oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı	58
Şekil 19: B Lenfosit alt gruplarının sayılarının yaşa göre durumunun DBK sınıflamasına göre dağılımı	59
Şekil 20: T Lenfosit alt gruplarının oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı	60
Şekil 21: T hücre alt gruplarının sayılarının yaşa göre durumunun DBK sınıflamasına göre dağılımı	60

Şekil 22: CD4 ve CD8 T lenfosit alt gruplarının oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı	61
Şekil 23: Ig seviyelerinin yaşa göre durumunun PKE gruplamasına göre dağılımı	62
Şekil 24: Hemogramın yaşa göre durumunun PKE gruplaması göre dağılımı	63
Şekil 25: Basit immünofenotiplemede hücre oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı	63
Şekil 26: B Lenfosit alt gruplarının oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı	64
Şekil 27: T hücre alt gruplarının oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı	65
Şekil 28: CD4 ve CD8 alt gruplarının oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı	65
Şekil 29: Hastalık başlangıcındaki semptom ve bulgular	74
Şekil 30: Hastalık seyrinde görülen semptom ve bulgular	74
Şekil 31: Hasta başına HPO terimi sayısı	75
Şekil 32: Terim azaltma öncesi ve sonrası hasta başına tespit edilen HPO terim sayısı dağılımı	76
Şekil 33: HPO terimleri birleştirme süreci	77
Şekil 34: HPO birleştirme sonrası ağırlıklı semptom dağılımları	85

1. GİRİŞ VE AMAÇ

İnflamatuvar Bağırsak Hastalıkları (İBH), gastrointestinal sistemde tekrarlayan inflamasyonlara sebep olan kronik bir hastalık grubunu ifade etmektedir. Klasik olarak İBH, Crohn Hastalığı (CH), Ülseratif Kolit (ÜK) ve sınıflandırılmayan (*undiferansiye*) formları içermektedir. Klasik İBH formları genellikle poligenik etyolojiye sahiptir ve toplumda yaygın bir şekilde görülmektedir. Son yıllarda mulifakoriyel yaygın İBH formlarının yanı sıra klinik olarak bunlara benzerlik göstermekle beraber etyolojilerinde tek gen bozukluğu yer alan bir grup İBH'nın varlığı anlaşılmıştır. Bu özel hastalık grubu, monogenik İBH (mİBH) olarak adlandırılmaktadır. Yeni Nesil Dizileme (*Next-Generation Sequencing*, NGS) teknikleri sayesinde, şimdiye kadar 100'den fazla mİBH tanımlanmıştır ¹.

Klasik İBH'ların aksine atipik seyirli olan İBH'lar hastalık bulgularının başlama zamanına göre infantil, çok erken başlangıçlı ve erken başlangıçlı olarak sınıflandırılabilir. Bu İBH tiplerinde hastalığın ilerlemesi son derece hızlıdır ve genellikle standart tedavilere direnç gösterir. Yapılan araştırmalar, mİBH etyolojilerinin bu grup hastalarda daha sık görüldüğünü göstermektedir ². Böyle durumlarda hastaların tanı ve tedavisinde birçok zorlukla karşılaşmaktadır. Bu özel grup içinde her bir çeşidin nadir olması nedeniyle, klinik gözlemler bazen sınırlı hasta verilerine dayanmaktadır. Hatta aynı genetik bozukluğa sahip bireyler arasında bile klinik bulgular, hastalık seyri ve tedavi yanıtları değişkenlik gösterebilir. Bu nedenle, mİBH hastalarıyla ilgili yayınlarda belirli standartlar oluşturma ihtiyacı vardır. Her bir olgu veya olgu serisinin farklı özelliklerinin yayınlanması ve belirli bir standart yaklaşımın eksikliği, literatürü gözden geçirirken zorluklara neden olmakta ve çok sayıda yayını içeren meta analiz gibi derlemelerin gerçekleştirilmesini imkansız hale getirmektedir. Son zamanlarda geliştirilen fenotip ontolojisi terimlerine dayalı algoritmalar insan hastalıkları ile farklı fenotipik özelliklerin ilişkisini ortaya koyarak olası genetik etyolojileri öngören yaklaşımlar sunmaktadır. Bu yaklaşımların amacı farklı klinik bulgulara sahip olgularda

hekimlerin ayırıcı tanı yapmasını kolaylaştırmaktır. İlerleyen yıllarda bu algoritmaların klinik mİBH'ya tanısal yaklaşımı kolaylaştıracağı beklenebilir. Ancak bu hedefe ulaşabilmek için öncelikle her bir hastalık grubunun sistematik bir yaklaşımla ele alınması ve farklı klinik tablolar ile genetik etyolojiler arasındaki ilişkilerin ayrıntılı bir biçimde değerlendirmesi gerekmektedir.

İBH fenotipi ile ortaya çıkan gen defektlerinin epitel hücreleri, fibroblastlar, endotel hücreler, fagositler ve lenfositler gibi pek çok farklı hücre gruplarında ifade edilmektedir. Örneğin, *ADAM17*, *IL10*, *IL10RA*, *IL10RB*, *GUCY2C*, *IL21*, *LRBA*, *TTC7A* ve *XIAP* genleri bu özellikleri taşımaktadır ³. Farklı hücrelerde görev almalarının yanı sıra mİBH ile ilişkili genlerin moleküler fonksiyonları incelendiğinde, birçok farklı moleküler yolağın etkilendiği görülmektedir. Bu farklılıklar, mİBH'nın tanısal yaklaşım ve tedavi arayışları açısından büyük önem taşımaktadır. mİBH ilişkili genlerin pekçoğu immün sistemin elemanlarında görev alırlar. Bu nedenle, bu genlere ait bozukluklar barsak bulguları yanı sıra immünolojik kusura da yol açar. Tanım itibarı ile de primer immün yetmezlik (PIY) veya diğer adıyla doğuştan bağışıklık kusurları (DBK, *Inborn Errors of Immunity*, IEI) olarak sınıflandırılırlar ⁴. Zaman zaman hastalar ağırlıklı olarak enfeksiyon ve barsak dışı organ bulguları ile kendini gösterir. İmmünoloji, enfeksiyon ve romatoloji gibi diğer disiplinlere başvuru olabilir. Bunun aksine hastalığın ilerleyen evrelerinde bağışıklık yetmezliği görülmesi de mümkündür. Hastalar, barsak enflamasyonuna yönelik tedavilerin yanı sıra immün modülatör tedaviler ve enfeksiyona karşı profilaksi gibi destekleyici ajanlarla da tedavi edilirler. Örneğin İntravenöz İmmunoglobulin (*Intravenous Immunoglobulin*, IVIg) ve profilaktik antibiyotik tedavileri sıklıkla uygulanır.

mİBH'na yönelik tedavi stratejileri halen araştırılmaktadır. Klasik İBH'nda kullanılan tedavi ajanlarının mİBH üzerindeki etkileri genellikle sınırlı kalmaktadır ⁵. Bununla birlikte, temel genetik bozukluğun tespit edilmesi durumunda, fonksiyonel olarak etkilenen moleküler yollara yönelik tedavi uygulamaları mümkün olabilmektedir. Uygulanabilecek tedavi yaklaşımları arasında hematopoetik kök hücre nakli, gen tedavisi, yolağa özgü farmakolojik ajanlar ve cerrahi müdahale

bulunmaktadır. Geniş bir tedavi seçeneği yelpazesinde, hastaya uygun olan etkili yöntemin seçilip uygulanması başarının temelini oluşturmaktadır.

Erken tanı, hem tedavi stratejilerini belirlemek hem de kalıcı doku hasarı ve mortaliteyi azaltmak için büyük önem taşır. mİBH ve klasik İBH arasındaki ortak ve ayırt edici klinik ve immünolojik bulguların tanımlanması erken tanı ve hedefe yönelik tedaviye olanak sağlayabilir. Örneğin, nötrofil disfonksiyonu ile ilişkili mİBH ile lenfosit eksikliği veya kompleman disregulasyonu nedeniyle ortaya çıkan mİBH'nın hücre veya yolak-özü klinik özellikleri farklı olabilir. Ya da sadece hematopoetik kök hücre nakli ile tedavi edilebilen mİBH ile konvansiyonel ajanlara yanıt veren mİBH arasında bazı ortak fenotipik özellikler olabilir. Ancak, literatürde bu grupları anlamlı bir şekilde birbirinden ayıran yöntemlerle ilgili çalışma sayısı sınırlıdır.

Genetik dizileme yöntemlerinin ilerlemesi ve genetik tanılarının artmasıyla birlikte fizyolojik yollar daha iyi anlaşılır hale gelmiştir. Bu durum, hastalık tanımlarının artmasına yol açmıştır. Ancak bu süreçte, genom çağı öncesinde yapılan hastalık sınıflandırması ve medikal terminolojinin yetersiz kaldığı görülmüştür. Bu sorunu çözmek amacıyla, yapay zeka algoritmaları kullanılarak hastalıkların daha iyi sınıflandırılması ve gen-fenotip-yolak ilişkilerinin daha net bir şekilde ortaya konulması için çalışmalar yapılmaktadır. İnsan Mendeliyen Kalıtları Çevrimiçi Veritabanı (*Online Mendelian Inheritance in Man*, OMIM), Mendeliyen hastalıkların kataloglandığı kaynaklar olarak sıkça kullanılmaktadır. Bununla birlikte, hala yeterli hızda güncelleme yapılamaması ve ilişkilendirme çalışmalarının sınırlı hasta örneklerine dayanması gibi kısıtlılıklar nedeniyle, farklı genetik altyapılara sahip mİBH hastalarının fenotipik özellikleri istenen doğruluk düzeyinde belirlenememektedir. Her biri nadir olan bu hastalıkların ayrı ayrı incelenerek klinik özelliklerinin daha iyi tanımlanması ve genotip-fenotip ilişkisinin doğru bir şekilde ortaya konulması gerekmektedir.

Çalışmamız, kliniğimizde yürütülen araştırma protokolleri kapsamında son 20 yıl boyunca toplanan olgu kohortları kullanılmıştır. Bu kohortları içinde çok geniş

çeşitlilikte genetik mutasyon mevcuttur. Hatta *CD55*⁶, *IL37*⁷, *PIK3R1*⁸, *ST6GALNAC1*⁹ ve *IRHOM2*¹⁰ mutasyonları ilk kez bu kohortlardaki hastalarda tanımlanarak literatüre eklenmiştir.

Bu tez kapsamında GİS bulgusu olan DBK hastaları ile immün yetmezlik bulgusu olan mİBH hastalarının klinik ve fenotipik özellikleri güncel fenotipik yaklaşımlar (İnsan Fenotip Ontolojisi, *Human Phenotype Ontology*, HPO¹¹) eşliğinde tekrar gözden geçirilmiştir. Farklı kohortlardan süzülerek bir arada değerlendirilen hastaların multisistemik özellikleri tek tek incelenmiş ve hastalarda gözlenen HPO terimlerinin sıklıkları tespit edilmiştir. Bu hastalıkların bazıları protein kaybettiren enteropati (PKE) tablosu ile ilişkili olup bu çalışmada PKE ile PKE dışı başvuruların özellikleri de ayrıca kıyaslanmıştır. Çok tabakalı bir yaklaşım benimsenmiştir:

- i. Hastaların klinik özellikler HPO terimleri kullanılarak sistematik bir şekilde her hasta için kaydedilmiştir.
- ii. İmmünolojik değerlendirme ve genetik incelemeler eşliğinde DBK sınıflaması yapılmıştır (şayet bu başlık altında değerlendirilebiliyorsa).
- iii. mİBH sınıflandırmasına göre kategorize edilmiştir.
- iv. PKE ve PKE dışı klinik başvuruların farklılıkları değerlendirilmiştir.

Bu başlıklar ve her bir alt başlık ilgili analizlerle detaylandırılmış ve katmanlar arası geçişler bir arada değerlendirilmiştir. Makine öğrenim yöntemleri kullanılarak hastalıkların birbirinden ayrışmasını en iyi sağlayan klinik özellikler tespit edilmeye çalışılmıştır. Literatürde, bu hastalık gruplarında HPO terimlerinin dağılımını, sıklığını ve tanı için önemliliğini gösteren başka bir çalışma bulunmamaktadır.

2. GENEL BİLGİLER

2.1. Protein Kaybettirici Enteropati

Protein Kaybettirici Enteropati (PKE), bağırsaklardan normalden fazla protein atılımıyla seyreden bir klinik durumdur. Bu durum, çoğu bağırsak hastalığıyla ilişkilendirilebilirken, bazı bağırsak dışı hastalıklar ve durumlarla da ortaya görülebilmektedir ¹². Protein atılımı sonucunda, 3. boşluklarda sıvı kaybı meydana gelir. Tedavi genellikle altta yatan hastalığın tedavisine yöneliktir, bu nedenle kesin tanı koymak önemlidir.

PKE'de, protein kaybı iki ana mekanizma ile gerçekleşir. İlk mekanizma, lenfatik sistem anomalileri sonucu lenf sıvısı kaybıdır ¹³. Mukoza ve submukoza alanında bulunan lenf damarlarının obstrüksiyonu nedeniyle genişlemesi (İntestinal Lenfanjektazi, İL) sonucunda ortaya çıkar. Genişleyen damarlar zamanla patlar ve protein açısından zengin lenf sıvısı kaybı oluşur. Lenfatik obstrüksiyon, venöz sistemdeki basınç artışı (sarkoidoz, İBH, lenfoma, Fontan operasyonu sonrası vb.) kaynaklı olabileceği gibi bazen nedeni belirlenemez.

PKE'nin ikinci mekanizması, bağırsak mukozasında hasara bağlı olarak geçirgenlik artışıdır. Bağırsak mukozasının hasar görmesi en sık İBH gibi durumlarla ilişkilidir ¹⁴. İBH aynı zamanda obstrüksiyon nedeniyle de PKE'ye yol açabilir. Bunun dışında Salmonella ¹⁵, Şigella ¹⁵ ve Giardia ¹⁶ gibi enfeksiyonla, Çölyak hastalığı ¹⁴, Hemolitik Üremik Sendrom (HÜS) ¹⁷ ve besin alerjilerinde ^{18,19} de PKE görülebilmektedir.

2.2. Primer İntestinal Lenfanjektazi

İL, lenfatik sistemi etkileyen sistemik bir hastalığın ya da tanımlanmış bir genetik mutasyonun sonucunda görülebildiği gibi hiçbir sistemik bulgu olmadan, izole olarak, da hastalarda bulunabilir ²⁰. Herhangi bir hastalığa bağlı olmadan gelişen tipi Primer İntestinal Lenfanjektazi (PİL) olarak adlandırılır ¹³.

PİL hastalarında kan albumin ve total protein seviyeleri düşük bulunmaktadır ²¹. IgG, IgA ve IgM seviyelerinde düşüklükler, B ve T lenfosit yetersizliklerini göstermektedir. Lenfosit alt gruplarından Naive CD4 ve Bellek CD8 T lenfosit düşüklüğü görülebilmektedir ^{19,22}. Bu özellikler PİL'nin immün yetersizlikler ile yakın ilişkide olduğunu göstermektedir ²³.

2.3. İnflamatuvar Bağırsak Hastalıkları

İBH, gastrointestinal sistemin kronik inflamasyonu ve immün aktivasyonu sonucu oluşan bir hastalıktır ²⁴. İBH'nın oluşumunda genetik faktörler ve çevresel koşullara bağlı mikrobiyota değişikliklerinin etkili olduğu düşünülmektedir ²⁵. İBH'nın insidansı ve prevalansı bölgeden bölgeye farklılık göstermektedir. Özellikle endüstrileşmiş ülkelerde İBH oranları daha yüksektir. Kuzey Amerika'da 1,5 milyon, Avrupa'da ise 2 milyon kişinin hastalıktan etkilendiği bilinmektedir ²⁶. Ülkemizde ise insidansı 100 binde 1,86 - 3,09 arasında değişirken, prevalansı ise 100 binde 2,42 - 21 arasında değişmektedir ²⁷.

İBH genelde 15-30 yaşları arasında görülür ²⁸, ancak vakaların %25 kadarı 20 yaşından önce bulgu verir ²⁹. Çok merkezli pedyatrik İBH kayıtlarına göre ortalama tanı yaşı 10,3'tür. Hastaların %15'i 6 yaşından önce, %48'i 6-12 yaş arası ve %37'si de 13-17 yaş aralığında tanı almaktadır ³⁰. Olguların %24-32'sini Ülseratif Kolit

(ÜK), %59-73'ünü Crohn Hastalığı (CH), %3-13'ünü ise sınıflandırılmayan İBH (sİBH) oluşturmaktadır ³¹.

İBH'nda, hem gastrointestinal sistem (GİS) hem de GİS dışı semptomlar görülebilir. Hastaların en yaygın GİS semptomları ishal, rektal kanama, karın ağrısı ve tenesmustur (dışkılama hissi olmasına rağmen dışkılamada zorlanma). GİS dışı tutulumda ise büyüme-gelişme geriliği, ateş, halsizlik, ağız ülserleri, çomak parmak, döküntü, üveit, sarılık ve/veya hepatomegali ve artrit görülebilir. Muayenede karın ağrısı ve/veya kitle hissi, fistül, fissür, veya *skin tag* (cilt üzerinde yumuşak cilt kitleleri) gibi bulgular tespit edilebilir ³². GİS dışı bulgular genellikle hastalığın aktivitesine bağlı olarak artış gösterir ³³. ÜK ile karşılaştırıldığında, CH daha sık GİS dışı tutulum gösterir ³⁴.

ÜK, tipik olarak rektumdan başlayarak proksima yönde kesintisiz bir şekilde ilerleyen ve mukozayı süreğen olarak tutan bir İBH türüdür. En sık başlangıç semptomları rektal kanama, ishal ve karın ağrısıdır ³⁵. ÜK, GİS dışı sistemleri etkileyebilir, ancak GİS içinde sadece kolonda ve terminal ileum tutulumu görülür.

Pediyatrik başlangıçlı ÜK, yetişkin başlangıçlı ÜK'ten bazı noktalarda ayrılır. Pediyatrik başlangıçlı olguların %60-70'inde pankolit görülmekteyken yetişkin başlangıçlı olguların sadece %20-30 kadarında pankolit görülür ³⁰. Proktit ise çocuklarda nadir olarak görülen bir bulgudur ^{36,37}. Bazı hastalarda bağırsak dışı bulgular, bağırsak semptomlarından daha önce çıkabilir ³⁸.

CH, gastrointestinal yolda tüm organlarda görülebilen bir İBH türüdür. En sık tutulan bölgeler terminal ileum, kolon ve perianal bölgedir ²⁴. İnce ve kalın bağırsağın birlikte tutulumu %59 oranında görülürken, sadece kolon tutulumu %28, sadece ince bağırsak tutulumu ise %16 oranında gözlenmektedir ³⁹.

CH, mukozayı tam katman halinde etkileyen bir özelliğe sahiptir. Biyopsi örneklerinde granülomlar görülmesi, CH için özgün bir bulgu olarak kabul edilir ⁴⁰. ÜK'in aksine, CH aralıklı tutulum gösterebilir ve tutulum gösteren bölgeler arasında

normal doku bulunabilir. Bu nedenle, endoskopik görüntüleme sırasında "kaldırım taşı manzarası" olarak adlandırılan bir görüntü oluşabilir.

CH olan hastalarda en sık görülen şikayetler kilo kaybı/büyüme geriliği, karın ağrısı ve ishaldir ²⁴. Bulguların ve tanının daha küçük yaşlarda ortaya çıktığı çocuklarda izole kolon tutulumu daha sık görülebilir ⁴¹.

Her ne kadar İBH olgularının büyük çoğunluğu ÜK ya da CH olarak tanımlansa da bazı hastalar bu tanımlama dışında kalmaktadır. Sınıflanamayan İBH olarak adlandırılan bu grup, klinik olarak İBH olduğu açık olan ancak kesin olarak bir gruba dahil edilemeyen hastalardan oluşmaktadır. Özellikle çocuklarda daha sık görülmekte, yaş ile ters orantılı dağılım göstermektedir. Bu hastalarda atipik semptomlar ve pankolit daha sıklıkla görülmektedir ⁴². Ailevi yatkınlığın daha yaygın görüldüğü bu grupta, genetik etyolojinin daha önemli olduğu düşünülmektedir ⁴³. Yapılan genetik çalışmalar, sınıflandırılmayan İBH'nin bir kısmının tek gen mutasyonuna bağlı olduğunu göstermiştir ⁴².

2.4. Erken ve Çok Erken Başlangıçlı İnflamatuvar Bağırsak Hastalıkları

İBH, etyolojiye bağlı olarak sınıflandırılabilceği gibi başlangıç yaşına göre de sınıflandırılabilir. 2005 yılında yayınlanan Montreal Sınıflaması, 17 yaşın altındaki tüm olguları A1 sınıfına dahil etmiş ve bunları "Pediatrik başlangıçlı İBH" olarak adlandırmıştır ⁴⁴. Daha sonraki çalışmalar, 10 yaşından küçük çocuklarda, büyüklere göre farklı fenotipik özelliklerin olduğunu göstermiştir. CH hastalarında ileal tutulumun daha az, pankolitin daha yaygın olduğu, *anti-Saccharomyces cerevisiae* antikorlarının daha düşük olduğu; ÜK hastalarında ise kolektomi ve biyolojik ajanların daha sık kullanım gerektiren durumlar olduğu gözlenmiştir. Bu nedenle, 2011 yılında Paris Sınıflaması adı verilen ve Montreal Sınıflamasını geliştiren bir sınıflandırma sistemi oluşturulmuştur ⁴⁵. Bu sınıflamaya göre, 10 yaş

altı olgular A1a olarak sınıflandırılırken, 10-17 yaş arası olanlar ise A1b olarak sınıflandırılmış ve A1a sınıfı "Erken başlangıçlı İBH (EBİBH)" olarak adlandırılmıştır.

EBİBH olarak sınıflandırılan hastaların daha ayrıntılı analizlerinde, özellikle 6 yaş ve altı grupta, 7-10 yaş grubuna kıyasla daha fazla ailevi yatkınlık ve tek gen mutasyonlarının daha sık görüldüğü gözlemlendi. Bu nedenle, bu özel grupta ayrı bir sınıflandırma yapılarak "Çok Erken Başlangıçlı İBH (ÇEBİBH)" terimi kullanılmaya başlandı ⁴⁶.

Paris Sınıflaması içerisinde üzerinde durulan bir diğer grup, özellikle 1. derece akrabalarında İBH öyküsü daha sık görülen ve 1 ay ile 2 yaş arasında bulgular ortaya çıkan İnfantil Başlangıçlı İBH (İBİBH) grubudur. Bu grup, diğer gruplara göre daha agresif seyirli ve tedaviye dirençli bir alt gruptur ³⁹. Lenfoid hücre infiltrasyonu ile seyreden otoimmün bozukluğun patogenezi ön planda olduğu bu grupta, mortalitenin de daha yüksek olduğu gözlenmektedir ⁴⁷. Ayrıca, PİY görülme olasılığı da diğer gruplara kıyasla daha yüksektir ⁴⁸.

2.5. Monogenik İnflamatuvar Bağırsak Hastalıkları

Genetik biliminin gelişmesi ile tıpta pek çok alanda olduğu gibi İBH alanında da değişime yol açtı. Özellikle atipik klinik görülen ve tedaviye dirençli olan İBH formlarının, monogenik etyoloji kaynaklı olduğu fark edildi ⁴⁹. Monogenik İBH (mİBH), klasik İBH'na kıyasla daha erken yaşta başlangıç gösteren, tedaviye dirençli ve bağırsak dışı bulguların daha belirgin olduğu bir İBH formudur ⁵⁰. Bu vakaların büyük çoğunluğu 2 yaş altında bulgular ortaya çıkarmıştır.

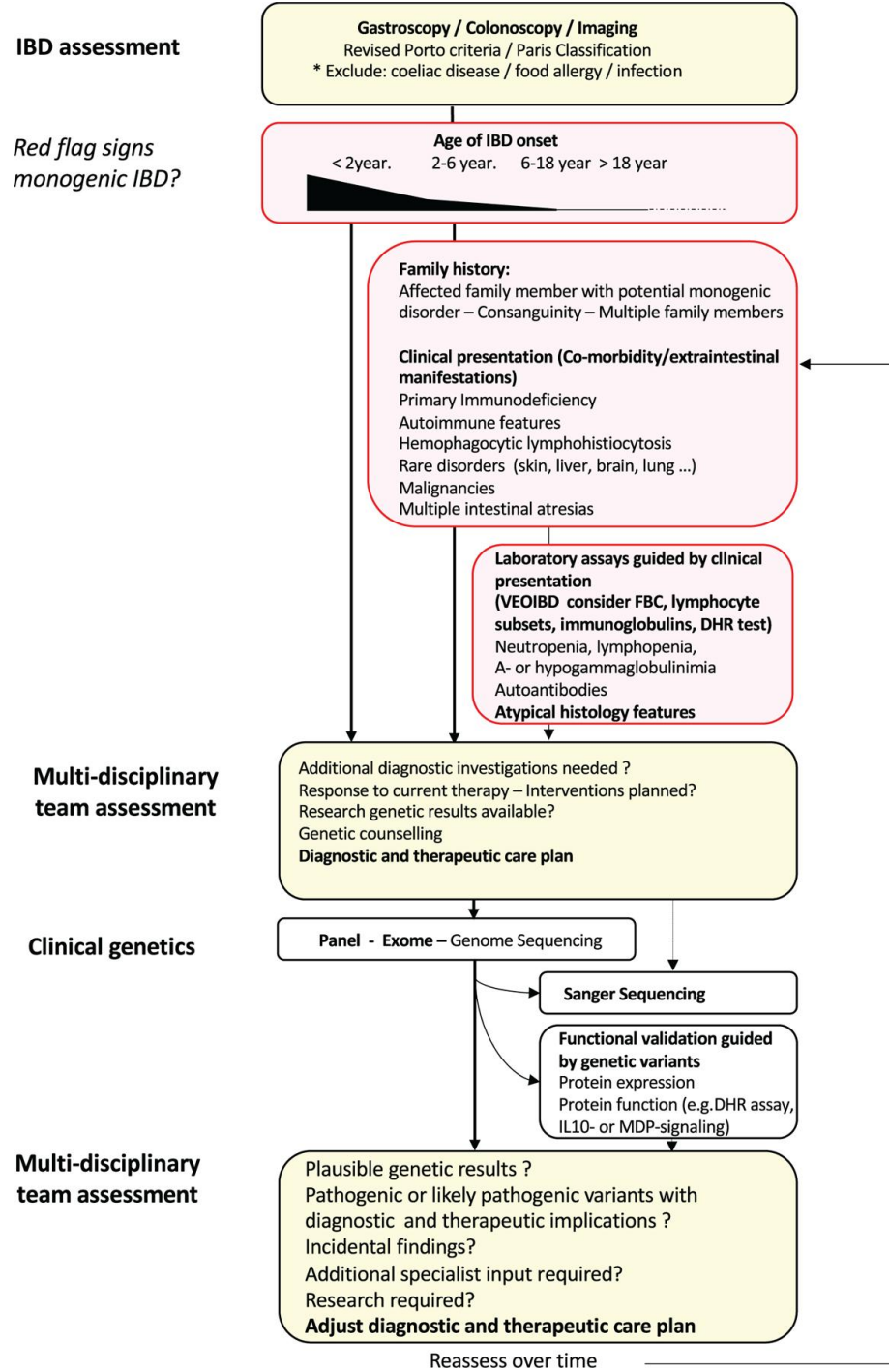
Bugüne kadar monogenik İBH ile ilişkili 102 gen tanımlanmıştır ¹ ve bu sayı gün geçtikçe artmaktadır. Literatürde en sık raporlanan mİBH genleri IL10RA, XIAP,

CYBB, IL10RB, LRBA, TTC7A dır. Genetik tanının netleşmesi, tedavi sürecinin yönetilmesi de kolaylaştırmaktadır. Etyolojiye yönelik ve hastaya odaklanan tedaviler (kemik iliği transplantasyonu, biyolojik ajan vb), klasik tedavilere kıyasla daha başarılı sonuçlar vermektedir. Ayrıca, ailelere genetik danışmanlık hizmeti sunulabilmektedir ⁴⁴.

mİBH'da bağırsak dışı tutulum sık görülen bir durumdur. 6 yaşından büyük olan olguların yarısından fazlasında hastalık, GİS dışı semptomlar ile başlamıştır ⁴⁴. Olgularında %76'sında GİS dışı tutulum gözlenmektedir. En sık görülen tutulumlar atipik enfeksiyon (%44,7), deri/saç/tırnak anomalileri (%38,4) ve otoimmünitedir (%21,6) ⁴⁴. Klasik İBH'da gözlenen perianal tutulum ve oral ülserler mİBH'da daha nadir görülür.

mİBH için genetik tanı dışında spesifik bir tanı yöntemi bulunmamaktadır. Hangi hastaların genetik testlere yönlendirileceğini belirlemek her zaman kolay olmayabilir. Bununla birlikte, 2 yaş altı olgular, akraba evliliği, ailede İBH öyküsü, büyüme-gelişme geriliği, sık enfeksiyonlar, endokrinopati, ağır perianal tutulum, Makrofaj Aktivasyon Sendromu veya Hemofagositik lenfhistiositoz, intestinal obstrüksiyon, cilt bulguları, tümör varlığı ve klasik tedavilere dirençli olgularda mİBH'dan şüphelenilmelidir ⁵¹.

mİBH düşünülen olgularda, hemogram, immunoglobulin düzeyleri, nötrofil aktivitesi ve lenfosit alt grup analizi gibi testler yapmak özellikle olası genetik mutasyonları ön görmeye yardımcı olabilir ⁵¹. mİBH şüphesi olan olgulara yaklaşım Şekil 1 de özetlenmiştir ⁵².



Şekil 1: mIBH düşünülen olgulara yaklaşım.

Şekil mIBH şüphelenilmesi için gerekli kriterleri, planlanması gerekli laboratuvar ve genetik tetkikleri anlatmaktadır. Uhlig ve arkadaşlarının yayınladığı “Clinical genomics for the diagnosis of monogenic forms of inflammatory bowel disease: A position paper from the paediatric IBD porto group of european society of paediatric gastroenterology, hepatology and nutrition.” isimli çalışmadan alınmıştır⁵².

2.6. İmmün Yetmezlikler

İmmün yetmezlik, bağışıklık sistemindeki doğal veya edinilmiş hücre grupları veya sinyal yolları etkilendiğinde ortaya çıkan bir hastalık grubudur. Bu durumda, etkilenen sistem ve etkilenme derecesine bağlı olarak geniş bir yelpazede bulgu ve semptomlar görülebilir. Hastalar genellikle tekrarlayan ve kronik enfeksiyonlarla başvururken, bazılarında büyüme-gelişme geriliği, ağır atopi, otoinflamatuar veya otoimmün bulgular başvuru şikayeti olabilir ^{53,54}. Hastaların çoğunluğu çocukluk döneminde bulgular gösterirken, hafif formlarda semptomsuz seyir erişkin yaşlara kadar devam edebilir ⁵⁵.

İmmün yetmezlikler, primer ve sekonder olmak üzere iki ana kategoriye ayrılabilir. Sekonder immün yetmezlikler, başlangıçta normal işleyen bağışıklık sisteminin sonradan enfeksiyonlar, ilaçlar, kanser gibi nedenlerle bozulmasıyla ortaya çıkar. Bu tür immün yetmezliklerin büyük çoğunluğunda, altta yatan neden düzeltildiğinde bağışıklık sistemi de iyileşir.

Primer immün yetmezlikler, genellikle monogenik mutasyonlar nedeniyle doğuştan gelen immün yetmezlik durumlarıdır. PİY’de hem hücresel hem de humoral bağışıklık sistemi, ayrı ayrı veya birlikte etkilenebilir. Bu hastalarda fırsatçı enfeksiyonlar sıklığı artabilir veya spesifik patojen gruplarına karşı yatkınlık gelişebilir ^{56,57}. Her bir genetik mutasyonun prevalansı tek başına düşük olabilir, ancak toplu olarak düşünüldüğünde nadir denilemeyecek kadar sıklıkla görülmektedir. Prevalansları 100.000 kişide 4.4 ile 83 arasında değişmektedir ⁵⁵.

Literatürde PİY ile ilişkili 500’den fazla genetik mutasyon vardır. Uluslararası İmmünoloji Dernekleri Birliği (*International Union of Immunological Societies - IUIS*) çatısında bir araya gelen uzmanlar bu mutasyonları birbirlerine olan benzerliklerine göre 10 başlık altında toplamıştır ⁴. Bu sınıflamanın özeti Tablo 1’de gösterilmiştir.

Görüldüğü gibi, PİY geniş bir dağılım gösteren bir hastalık grubudur. PİY tedavisinde IVIg, profilaktik antibiyotik, antifungal ve/veya antiviral kullanımı, kök hücre nakli, timus nakli, biyolojik ajanlar, enzim replasmanı ve gen tedavileri gibi yöntemler kullanılabilir. Hastalarda akciğer veya GİS tutulumu, maligniteler, lenfoproliferatif hastalıklar, otoimmünite ve endokrinopatiler gibi durumlar ortaya çıkabilir. Hastalar düzenli olarak görme, işitme ve diş muayenelerine tabi tutulmalıdır. Doğru ve hızlı bir şekilde doğru tanıya ulaşmak, hem primer hastalığın tedavisi hem de eşlik eden komorbiditelerin takibi açısından büyük önem taşımaktadır.

Jeffrey Modell Vakfı tarafından geliştirilen uyarıcı bulgular listesi, PİY yönünden araştırılması gereken hastaların seçiminde kullanılan bir rehberdir ⁵⁸. Tablo 2’de sıralanan bu listede belirtilen 10 maddenin hastada 2 veya daha fazlasının görülmesi durumunda, hastanın uzman bir kliniğe yönlendirilmesi önerilmektedir. Bu şekilde, potansiyel PİY vakalarının hızlı bir şekilde tanı ve uygun tedaviye ulaşması hedeflenmektedir. Uyarıcı bulgular listesi, PİY’in erken tanısı için önemli bir araç olarak kullanılmaktadır.

PİY düşünülen hastaların değerlendirilmesi önemli bir aşamadır ve dikkatli bir öykü alımı ve fizik muayene ile başlamalıdır. Hastanın semptomların başlama yaşı, hangi belirtilerin öne çıktığı, göbek kordonu düşme zamanı, akraba evliliği durumu ve ailede PİY tanısı konmuş kişilerin varlığı gibi faktörler detaylı bir şekilde değerlendirilmelidir ⁵⁹. Fizik muayene sırasında, hastanın boy-kilo persentilleri, BCG aşısı izi, daha önce geçirilen enfeksiyonlardan kaynaklanan izler, cilt bulguları, lenf nodları ve karın muayenesi gibi noktalar özenle incelenmelidir. Ayrıca, dismorfik vücut bulguları da kontrol edilmelidir ⁵⁹. Bu adımların uygulanması, doğru tanıya ulaşmak ve PİY olgularını tespit etmek açısından oldukça önemlidir. Öykü ve fizik muayene sonuçları, hastanın daha ileri değerlendirme için uygun bir uzman kliniğe yönlendirilmesi konusunda klinik kararın temelini oluşturacaktır.

Tablo 1: DBK Sınıflamasının özeti (Mayıs 2022)

Sınıf	DBK Hastalık Sınıfı	Gen Sayısı	Karakteristik Özellikleri	Örnekler
I	Hücrel ve Humoral immün yetmezlikler	66	- Tekrarlayan viral, fungal ve bakteriyel enfeksiyonlar - Otoimmünite	- SCID - DOCK8 eksikliği - CD40L eksikliği - MHC I/II eksikliği
II	Sendromik kombine immün yetmezlikler	69	- Tekrarlayan viral, fungal ve bakteriyel enfeksiyonlar - Otoimmünite - İmmün olmayan özellikler - Dismorfi	- Hipe-IgE (STAT3) - Wiskott-Aldrich sendromu - Kıkırdak-saç hipoplazisi 22q11.2 delesyon sendromu - NEMO eksikliği
III	Antikor defekleri	45	Tekrarlayan bakteriyel sinüpolmoner enfeksiyonlar	- X-linked (BTK) and otozomal resesif agammaglobulinemi - Otozomal resesif hiper-IgM - Selektif IgA eksikliği
IV	İmmün düzensizlik hastalıkları	52	- Lenfoproliferasyon - Otoimmünite - HLH	- Chediak-Higashi sendromu - Perforin eksikliği - ALPS - STAT3 GOF - CTLA4 eksikliği - LRBA eksikliği - IL-10 eksikliği
V	Fagosit hastalıkları	42	- Cilt, akciğer ve lenf nodlarının ağır bakteriyel enfeksiyonları - İnflamatuvar bağırsak hastalığı	- Kronik granülomatöz hastalık - Lökosit adezyon defektleri - GATA2 eksikliği - Konjenital nötropeniler
VI	Doğal bağışıklık yetmezliği	74	- Genelde ateş görülmeyen invaziv bakteriyel enfeksiyonlar (sepsis, menenjit) - Tekrarlayan fungal, mikobakteriyel ve viral enfeksiyonlar	- IRAK4 ve MyD88 eksikliği - STAT1 GOF - İnterferon gama reseptör eksikliği - IRF7 eksikliği - TLR3 eksikliği
VII	Otoinflamatuvar bozukluklar	56	- Tekrarlayan ateş, döküntü, artralji / artrit - Amiloidoz - İnflamatuvar bağırsak hastalıkları	- Ailevi Akdeniz Ateşi - Mevalonat kinaz eksikliği - Muckle-Wells sendromu - NLRP1 eksikliği
VIII	Kompleman eksiklikleri	36	- Dissemine Naisseria enfeksiyonları - Tekrarlayan ateşli enfeksiyonlar - SLE - Atipik HÜS	- Kompleman komponent eksiklikleri (C1 to C9) - Properdin eksikliği - Faktör D eksikliği - Faktör H eksikliği
IX	Kemik iliği yetmezlikleri	44	- Makrositer anemi ve diğer sitopeniler - Kemik, tırnak, saç ve/veya cilt anomalileri	- Fanconi anemisi - Dyskeratosis congenita
X	Primer immün yetmezlikleri fenokopyaları	15	Değişken fenotipler	- ALPS-SFAS (<i>TNFRSF6 somatik mutasyonu</i>) - IL-17 ya da IL-22 otoantikorları nedeni - Mukokutanöz kandidiyazis - IL-6 otoantikorlarına bağlı Mikobakteriyel enfeksiyonlar. - C1 inhibitör otoantikorlarına bağlı edinilmiş anjiödem - IFN-alpha ve IFN-omega otoantikorlarına bağlı ağır, yaşamı tehdit eden SARS-CoV2 enfeksiyonları

*IUIS sınıflamasına göre DBK sınıfları ve genel özellikleri. Tablo, UpToDate'ten alınmıştır*⁶⁰.

Tablo 2: PİY için uyarıcı bulgular

1. Yılda 4 ya da daha fazla yeni kulak enfeksiyonu.
2. Yılda 2 ya da daha fazla ağır sinüzit enfeksiyonu.
3. 2 ay ya da daha uzun antibiyotik kullanımına rağmen tedavi yanıtı olmaması.
4. Yılda 2 den daha fazla zatürre.
5. Büyüme ve gelişme geriliği.
6. Tekrarlayan cilt ya da organ apseleri.
7. Uzun süreli ağızda pamukçuk ya da ciltte mantar enfeksiyonu.
8. İntravenöz antibiyotik tedavisi ihtiyacı.
9. Septisemi dahil 2 ya da daha fazla derin doku enfeksiyonu.
10. Ailede PİY tanılı birey olması.

Jeffrey Modell Vakfı'nın web sitesinden alınmıştır ⁵⁸.

PİY tanısı için kullanılabilecek laboratuvar tetkikleri geniş bir yelpazeye sahiptir. Ancak tetkiklerin istenmesi ve değerlendirilmesi sırasında hastanın hikayesi ve fizik muayene bulguları dikkate alınmalıdır. Yaşa uygun normallere göre değerlendirme yapılmalıdır ve tetkik sonuçlarındaki değişkenliklerin aktif enfeksiyon veya hastalık dönemlerinde ortaya çıkabileceği unutulmamalıdır. Tetkik istemlerine sırasıyla hemogram, immünoglobulinler, aşılanan etkenlere karşı antikor yanıtları, polisakkarit etkenlere antikor bakılamıyorsa kan grubu ve izohemaglutinin ve akım sitometrisinde lenfosit alt grupları istemi yapılması önerilmektedir ⁵⁹. Bu tetkiklerin istenmesi, hastanın klinik özelliklerine ve şüphelenilen PİY alt tipine göre değişiklik gösterebilir. Hastanın kapsamlı bir değerlendirme ve uygun laboratuvar tetkikleri ile desteklenen tanısı, tedavi planlaması ve takibi için önemlidir.

Akım sitometrisi, lenfosit alt gruplarının belirlenmesinde kullanılan bir yöntemdir ve PİY etyolojisi hakkında önemli bilgiler sağlayabilir. Hastanın lenfosit alt gruplarının dağılımı, spesifik PİY alt tipleriyle ilişkili olabilir ve tanı sürecinde değerli bir belirteç olarak kullanılabilir. Tablo 1’de gruplara yakın bir lenfosit alt grupları dağılımı gösteren hastalar için, tüm genom veya egzom çalışmaları gibi daha

maliyetli yöntemler yerine, hedeflenmiş gen paneli testleri kullanmak daha maliyet etkin olabilir. Bu panel testleri, spesifik genlerin veya mutasyonların analizini hedefler ve PİY etyolojisini aydınlatmada yardımcı olabilir. Bazı durumlarda, örneğin ADA eksikliği gibi belirli PİY alt tiplerinde, enzim çalışmaları bile tanı koymada kullanılabilir ve tedavi sürecini başlatmaya yardımcı olabilir. Bu nedenle, lenfosit alt gruplarının dağılımıyla birlikte spesifik testlerin kullanılması, PİY tanısının doğrulanmasında ve uygun tedavi planının belirlenmesinde önemlidir.

Ancak, tanı sürecinde kullanılacak testlerin seçimi ve yorumlanması, deneyimli uzmanlar tarafından yapılmalıdır. Her hasta için özelleştirilmiş bir yaklaşım gereklidir ve kapsamlı bir klinik değerlendirme ile birlikte tetkiklerin seçimi ve yorumlanması önemlidir. PİY araştırmasında, öykü, fizik muayene ve laboratuvar bulgularının seçimi için belirli kriterler belirlenmiş olsa da, hangi özelliklerin hastalıkları ayırt etme konusunda daha önemli olduğuna dair objektif bilimsel veriler mevcut değildir. Bu konuda, makine öğrenimi modellerinin kullanılması, önemli bulguların seçimi ve bu bulgulara dayalı olarak hastalık tahminlerinin yapılmasını mümkün kılabilir.

2.7. İmmun Yetmezlikler ve Monogenik İnflamatuvar Bağırsak Hastalıkları

GİS mukozasında immünite, üç ana bileşenden oluşmaktadır. İlk bileşen, etkili bir mukozal bariyer bulunmaktadır. Bu bariyer, patojenlerin doku geçişini engeller ve antijenlerin, antimikrobiyal proteinlerin ve büyüme faktörlerinin salgılanmasını sağlar. İkinci bileşen doğal bağışıklık sistemidir. Bu sistemde, miyelositer hücreler, Doğal Öldürücü (*Natural Killer* - NK) lenfositler ve bazı lenfoid hücre grupları yer almaktadır. Bu mekanizmanın temel işlevi, patojenlerin tanımlanmasıdır. Üçüncü bileşen ise adaptif bağışıklıktır. Antijen sunucu hücreler aracılığıyla patojen proteinlerine karşı B ve T hücreleri tarafından antikor üretimi sağlanır. Bu sayede spesifik bir savunma yanıtı oluşturulabilir ⁶¹. Mukozal bariyerdeki immün

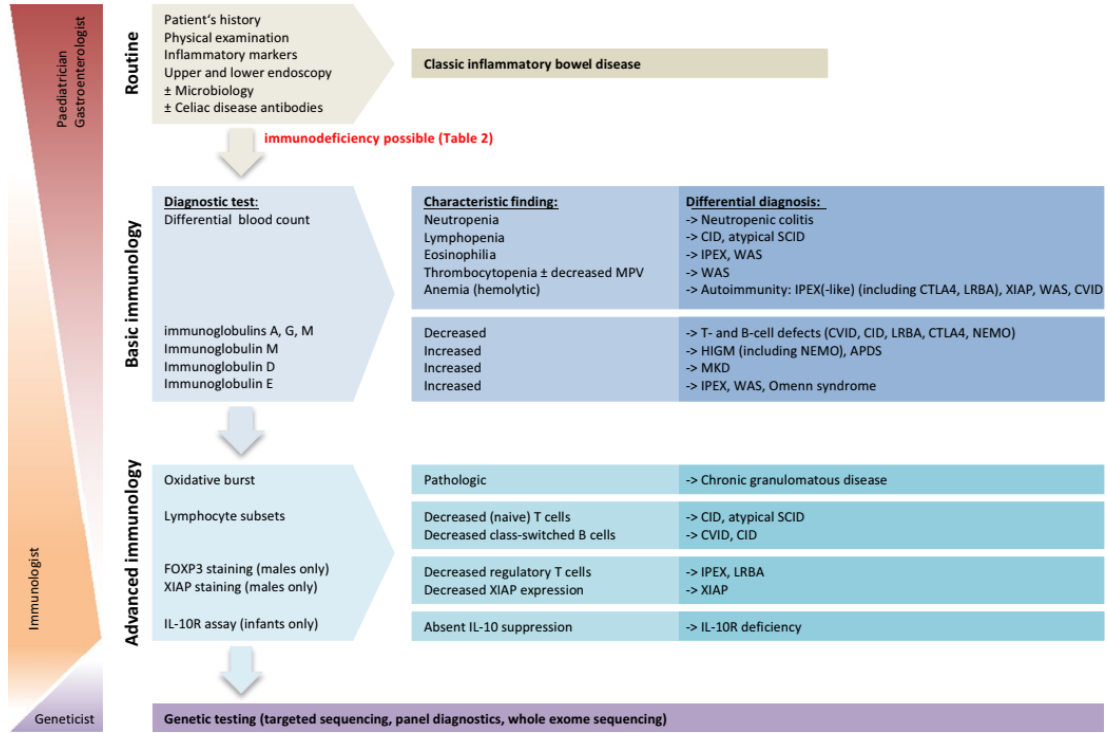
disregülasyonun İBH'na nasıl yol açabildiği konusunda birden fazla teori bulunmaktadır.

Epitel bariyerdeki düzensizlik: Bağırsak epitel hücreleri, antijen sunucu hücrelerle benzer proteinleri sentezleyebilme ve kullanabilme yeteneğine sahiptir ⁶². Bu hücrelerin anormal antijen sunumu yapması ve lenfositlerle etkileşime girmesi, İBH'nın bir nedeni olabilir ⁶³.

Bağıışıklık hücrelerinde düzensizlik: İBH hastalarından alınan biyopsi örneklerinde lamina propriada bazı sitokinlerin artışının olduğu gözlenmiştir ^{64,65}. İntestinal makrofajlardan salgılanan bu sitokinlerin NK hücreleri, T hücreleri, doğal lenfoid hücreler ve B hücrelerinin sayısını ve etkinliğini arttırdığı; bunun da İBH ile ilişkili olduğunu çalışmalarda göstermiştir ⁶⁶⁻⁷².

Düzensiz mediatör salgılanması: Spesifik bir bağıışıklık hücre grubunun normalden daha aktif çalışması, İnterferon gamma, Tümör Nekroz Faktör alfa, İnterlökin 4, İnterlökin 5, İnterlökin 13 gibi proinflamatuvar sitokinlerin artışına, buna bağlı olarak da İBH gelişimine yol açabilir. ⁷³⁻⁷⁸. Bu yolların düzensiz olduğu hastalarda, yolların bloke edilmesiyle tedavi sağlanabilir.

Genom boyu ilişkilendirme çalışmaları (*Genome-wide Association Studies*, GWAS), tüm egzom sekanslama (*Whole Exome Sequencing* - WES) ve tüm genom sekanslama (*Whole Genome Sequencing* - WGS) gibi teknikler kullanılarak PİY ile ilişkili 500 den fazla genetik bölge tespit edilmiştir ⁷⁹. Bu bölgelerin bazıları otoimmün (İBH, multiple skleroz, romatoid artrit vs) hastalıklarla da ilişkili bölgelerdir. Diğer taraftan ise mİBH nedeni olduğu tanımlanan 102 genin ¹ 75 tanesi DBK sınıflamasına ⁴ göre DBK olarak da tanımlanmaktadır. Her iki hastalık arasında yakın ilişki olduğu net olarak gözlenmektedir.



Şekil 2: PİY şüpheli İBH olgularına yaklaşım.

MPV, mean platelet volume (ortalama platelet hacmi); SCID, severe combined immunodeficiency (ağır kombine immün yetmezlik); CID, combined immunodeficiency (kombine immün yetmezlik); CVID, common variable immunodeficiency (yaygın değişken immün yetmezlik). Şekil Tegtmeier D. ve arkadaşlarının “Inflammatory bowel disease caused by primary immunodeficiencies-Clinical presentations, review of literature, and proposal of a rational diagnostic algorithm” adlı çalışmasından alınmıştır⁸⁰.

Histopatolojik yöntemler mİBH tanısına yardımcı olsa da PİY birlikteliğini kanıtlamakta yeterli değildir. Erken başlangıçlı hastalık, akraba evliliği, parenteral nutrisyon ihtiyacı, sık tekrarlayan hastalık, atipik histopatolojik bulgular, deri / saç / tırnak anomalileri, otoimmünite, tekrarlayan enfeksiyonlar, nötropeni / lenfopeni, anormal Ig değerleri olan hastalarda PİY araştırılması önerilmektedir⁸⁰. Şekil 2’de İBH olup PİY düşünülen hastalara yaklaşım gözetlenmiştir.

Genel olarak PİY ve mİB hastalık grupları birbiri ile içiçe geçmiş tıbbi durumlardır. Her iki durumda da benzer alarm semptomları mevcuttur ve bu

semptomlar klinisyenleri daha ileri arařtırmalar ve genetik testlere yönlendirmektedir. Günümüzde henüz her mİBH kuřkulu hastadan WGS çalışılması mümkün değildir. Hedeflenmiş gen panelleri daha yaygın olarak kullanılabilir. Ancak panel seçimi için hastaların klinik fenotipleri ve basit laboratuvar tetkikleri ile ayırım yapacak yöntemler geliştirilmesi gerekmektedir.

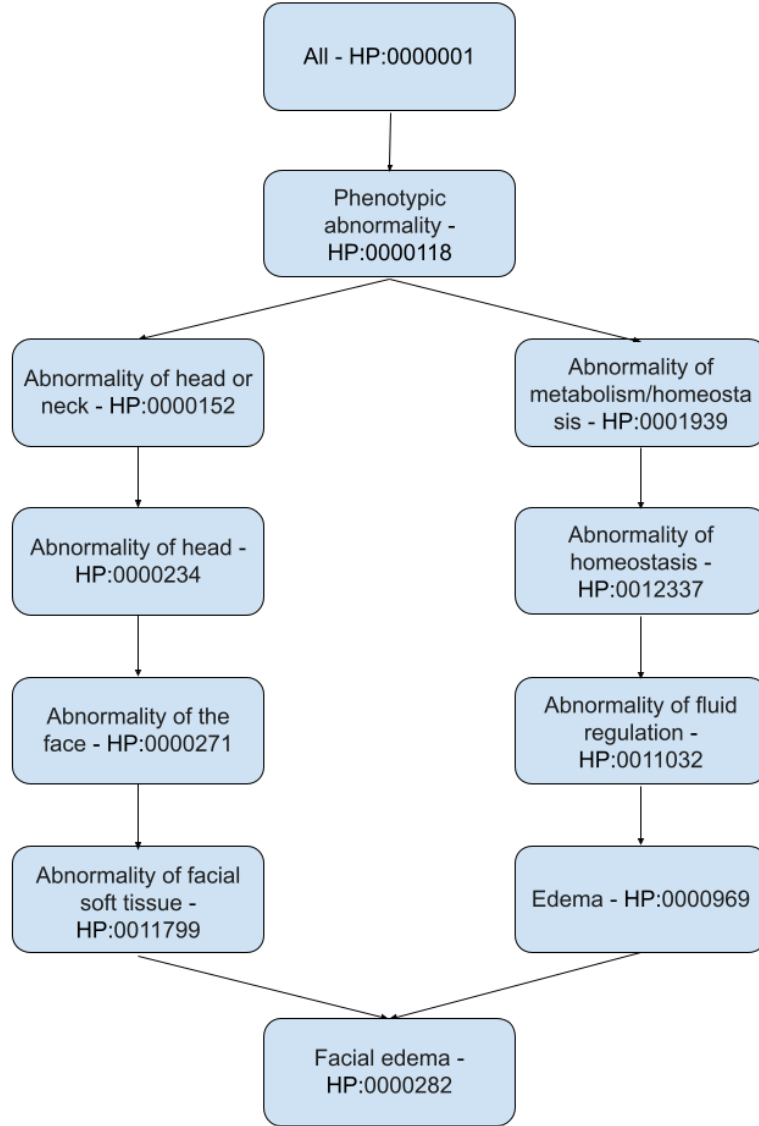
2.8. İnsan Fenotip Ontolojisi (*Human Phenotype Ontology, HPO*)

Ontoloji, Eski Yunanca’da “var olan” anlamına gelen “*ontos*” ve “mantıksal söylem” anlamına gelen “*logia*” kelimelerinin birleştirilmesiyle oluşturulan ve “varlığın çalışması/bilimi” anlamına gelen bir terimdir ⁸¹. Ontolojiler, felsefenin metafizik dalından çıkan, “ne vardır?” ya da “gerçekliğin doğası nedir?” gibi sorulara cevap arayan tartışmalardır. 1970’li yıllarda yapay zeka üzerine çalışan bilim insanları, güçlü yapay zekalar üretebilmek için bilgi mühendisliğinin kilit olduğunu fark etmiştir. 1980’lerde ontoloji, “bilgi tabanlı sistemler” için kullanılan bir terim iken 1990’larda anlamsal ağlar veya taksonomik sınıflamaları işaret eden bir terim olarak kullanılmaya başlanmıştır ⁸².

Bilim insanları, genlerin ve gen ürünlerinin anlamsal olarak sınıflandığı bir sistemine ihtiyaç duymuştur. Bu amaçla, 1998 yılında Ashburner ve arkadaşları tarafından *Drosophila melanogaster* (meyve sineği), *Mus musculus* (fare) ve *Saccharomyces cerevisiae* (bira mayası) genlerini içeren ilk Gen Ontolojisi’ni (*Gene Ontology, GO*) yayınlanmıştır. GO sayesinde ilk defa bilgisayarlar tarafından okunabilen, birleştirilmiş, açıklamalı bir terim kütüphanesi oluşturulmuş durumdadır⁸³.

GO çalışmaları ile başlayan sınıflama çalışmaları ile 2001 yılında Açık Biyolojik ve Biyomedikal Ontolojiler Atölyesi’nin (*The Open Biological and Biomedical Ontologies Foundry*, daha önceden *Open Biomedical Ontologies Foundry, OBO*)

kurulması sağlanmıştır ⁸⁴. Bu grupta halen devam eden çalışmalar, biyoloji ve tıp alanında uyumlu referans ontolojilerin oluşturulması için gerekli araçlar ve prensipleri ortaya koymaktadır. Aynı zamanda, yeni ontolojilerin geliştirilmesi ve mevcut ontolojilerin sürdürülmesi konularında da çalışmalar yürütülmektedir. OBO, şu an itibarıyla erişilebilir 186 adet ontolojiyi bünyesinde barındırmaktadır.



Şekil 3: HPO ağacından bir kesit

Şekilde görüldüğü gibi “Facial Edema” için 2 farklı yolak bulunmaktadır. Her iki yolak da 7 basamaktan oluşmaktadır.

GO'nun başarılı veri analizleri ve büyük veri setleri üzerinde çalışabilme yeteneğini gözlemleyen bilim insanları, insan fenotipik özelliklerinin ontolojisinin de verimli ve kullanışlı olabileceğini öngörmüştür. Bu nedenle, 2008 yılında Robinson ve arkadaşları tarafından, OMIM⁸⁵ verilerinden hareketle, ilk İnsan Fenotip Ontolojisi'ni (*Human Phenotype Ontology*, HPO) geliştirilmiştir⁸⁶.

HPO, diğer OBO ontolojilerinde olduğu gibi, doğrudan asiklik ağ şeklinde düzenlenmiş bir yapıya sahiptir. Her bir fenotipik terim (düğüm), kendisinden daha genel ve de varsa daha detaylı bir başka fenotipik terimle ilişkilidir (kenar). Tüm fenotipik özellikler, genel tanımlara doğru birleşerek “*All - HP:0000001*” terimi altında toplanır. Daha genel tanımlı düğümler ile daha detaylı tanımlamalar içeren düğümler her zaman hiyerarşik bir düzen ile bağlantılı olarak gösterilir. Bu nedenle, döngüsel olmayan (asiklik) bir düğüm haritası oluşur. Şekil 3'te görüldüğü üzere aynı alt düğümü gösteren birden fazla yolak olabilmektedir. Bu yollar birbiri ile aynı ya da farklı sayıda basamaklara sahip olabilmektedir.

Fenotipik özelliklerin tanımlanması kadar önemli olan bir başka konu da hastaların fenotipik özelliklerinin HPO terimlerine nasıl çevrileceğidir. Bu konuda uygulama birliği sağlamak için dört kural belirlenmiştir⁸⁷:

1. Var olan en spesifik HPO terimini kullanmak: HPO listesi 2 ayda bir güncellenmekte olup çalışma yapıldığı sırada en güncel olan liste ele alınmalıdır. Örneğin hastada sakküler morfolojide arkus aorta anevrizması var ise “*Aortic aneurysm - HP:0004942*” yerine, ondan daha detaylı bir tanım olan “*Saccular aortic arch aneurysm - HP:0031647*” terimi seçilmelidir. Aranılan terimin daha detaylı bir versiyonu bulunamadığı koşullarda eksik kalan terimler uygun kanallar aracılığıyla bildirilmelidir.
2. Önemli ve ilgili HPO terimlerini seçmek: Genel olarak hasta için önemli olan fenotipik özelliklerin seçilmesi, hastada geçici olarak görülen (araya giren enfeksiyon ya da hastalığın atak döneminde sekonder bozulan bulgular gibi) fenotipik özelliklerin dahil edilmemesi önerilmektedir. Burada “önemli”

olarak ifade edilen özelliklerin seçimi için kesin kriterler bulunmamaktadır. Özellikle çalışma içine dahil edilen probandlar arasında farklılık gösterebilir. Kesin karar çalışmaları yürüten klinisyenler tarafından verilmelidir.

3. Normal olan özelliklerin eklenmesi: Hastada normal bulunan bir özellik var ise bu özellik belirtilmelidir. Normal özellikler belirtilirken, patolojik özelliklerin aksine, mümkün olan en geniş tanımlı HPO terimi seçilmelidir. Örneğin detaylı Hepatobilier US görüntülemesi ile anatomik bozukluk olmadığı tespit edilen hastaya “*NOT Abnormal liver morphology* - HP:0410042” terimi tanımlanmalıdır.
4. Yeterli sayıda terim seçilmesi: Optimal sayı konusunda genel bir kural yoktur. Literatürde 5 adet iyi seçilmiş HPO terimi ile başarılı fenotip güdümlü egzom analizi yapılabildiği gösterilmiştir ⁸⁸. Bu nedenle her hasta için en az 5 HPO terimi tanımlanması önerilmektedir. Eğer daha az fenotipik özellik tespit edilebilirse, olmayan fenotipik özellikler de eklenerek sayı 5’e tamamlanmalıdır.

HPO'nun kullanımıyla hastalıkların ve genetik mutasyonların öngörülmesi mümkündür. Genetik mutasyonların HPO terimleriyle taranması ve benign mutasyonların elemesi, klinisyenlerin tanı koyma sürecini hızlandırır ve kolaylaştırır. Bu amaçla *Exomiser* ⁸⁹, *Phenomizer* ⁹⁰, *Orphanet* ⁹¹ gibi programlar geliştirilmiş ve araştırmacıların kullanımına sunulmuştur.

Exomiser, hastaya ait ham genom verisini, hastanın HPO terimleri ile kodlanmış fenotip verisi ile karşılaştırır. Hastanın mutasyonlarını fenotipi ile uyumuna göre eler ve en olası mutasyonları verir. Böylelikle klinisyenler sadece şüpheli genleri inceleyerek zamandan tasarruf eder. *Phenomizer* ve *Orphanet* ise verilen HPO terimlerine uyumlu olan tüm genetik mutasyonları sıralar. Klinisyenler bu araçlardan edindikleri öngörü ile seçilmiş genlere yönelik tetkik isteminde bulunabilir.

Literatürde ve klinik uygulamalarda, HPO aracılığıyla genetik tanıya yönelik birçok örnek bulunmaktadır. Bu örnekler genellikle OMIM veritabanıyla ilişkilendirilir. OMIM, genetik mutasyonların fenotipik özelliklerini listeler. Ancak hangi fenotipin hangi mutasyonda baskın olduğuna dair bilgi sunulmamaktadır. Bu nedenle, OMIM tabanlı taramaların en zayıf yönü, genotip-fenotip ilişkilerini içeren filtrelemelerin yapılmamasıdır. Çalışmamızda, gerçek vakalar üzerinde bu ilişkilerin gösterilmesi ve bu ilişkilerin makine öğrenme teknikleriyle hastalık sınıflandırması yapabilme başarısı incelenmiştir.

2.9. Yapay Zeka ve Makine Öğrenimi

Yapay zeka, makinelerin insanlar veya diğer hayvanlar gibi bilgiyi algılama, sentezleme ve çıkarımda bulunma yeteneğini ifade eder ⁹². Tarih boyunca, yapay zeka örneklerine dair birçok kaynak bulunmaktadır. Antik Yunan mitlerinde yer alan *Talos* adlı insansı savaş makinesi, tarihte karşılaşılan ilk yapay zeka örneği olarak kabul edilir. İnsanlar, çevreden gelen bilgileri değerlendirebilen ve bu bilgilere dayanarak sonuçlar üretebilen "akıllı" makinelerin üretimi konusunda uzun bir süredir çaba sarf etmişlerdir. Modern bilgisayarların icadından önceki dönemde, yapay zeka denemeleri su saatleriyle başlamış ve hareketli tanrı heykelleri, “*automata*”lar ve Mekanik Türk ⁹³ gibi analog sistemlerle devam etmiştir.

Onyedinci yüzyıllara kadar üretilen yapay zekaların ortak özelliği, sadece tek bir görevi yerine getirebilmeleridir. Bu tür yapay zekaların görevlerinde değişiklik yapmak veya tamamen farklı bir görevi yerine getirmek için makinenin kısmen veya tamamen değiştirilmesi gerekmektedir. Makinelerin çok amaçlı hale getirilmesi için insanlarla etkileşim kurma ihtiyacı fark edilmiştir. Bu etkileşimin matematiksel temelleri 17. ve 18. yüzyıllarda oluşmaya başlamıştır ⁹⁴. Bu çalışmaların sonucunda, 19. yüzyılda yeniden programlanabilir ilk otomatik dokuma tezgahı geliştirilmiştir.

Bu makine, delikli kartlar aracılığıyla istenilen desenleri kusursuz bir şekilde tekrar tekrar üretebilmektedir.

Yirminci yüzyılın başında Alan Turing tarafından ortaya atılan "hesaplama teorisi" ⁹⁵, matematiksel modeller aracılığıyla karmaşık problemlerin çözülebileceğini göstermiştir. 1943'de yapay nöral ağlar tanımlanmış ve 1956 yılında yapay zeka akademik bir disiplin haline gelmiştir. 1960-1980 arasında birçok yapay zeka teoremi geliştirilmiştir, ancak o dönemde bilgisayarların işlem kapasiteleri yeterli olmadığından anlamlı sonuçlar elde etmek mümkün olmamıştır. 1980'lerden sonra bilgisayar teknolojisi yeterince olgunlaşmış ve kullanılabilir yapay zeka ürünleri ortaya çıkmıştır. Günümüzde, görüntüyü anlama ve işleme, konuşma tanıma, dil işleme gibi birçok farklı yapay zeka uygulaması bulunmaktadır.

Yapay zeka, kapsamlı bir beceri yelpazesi içeren bir tanımdır. Bu tanım içerisinde, önceden programlanmış basit otomasyon sistemlerinden karmaşık görüntü ve ses işleme tekniklerine kadar çeşitli yetenekler bulunmaktadır. Yapay zekanın, girdiler ve çıktılar sağlandıktan sonra kendi kendini programlayabilen bir hâlde ise bu duruma "makine öğrenimi" adı verilmektedir ⁹⁶. Makine öğrenimi, denetimli öğrenme, denetimsiz öğrenme ve takviyeli öğrenme olmak üzere üç temel yaklaşımı içermektedir.

Denetimli öğrenme (*supervised learning*) olarak adlandırılan yöntemde, belirlenen bir ilişkilendirme yöntemi (algoritma), önceden tanımlanmış bir girdi ve çıktı setiyle eğitilir. Eğitim sürecinin sonucunda, eğitim setindeki verilere uygun ilişkileri öğrenen bir model elde edilir. Bu eğitilmiş model, eğitim setindeki veri ile aynı yapıya sahip yeni verileri, eğitim setindeki verilerin ilişkilerine dayanarak tahmin eder ⁹⁶. Denetimli öğrenme yöntemi genellikle sınıflandırma sistemlerinde kullanılır.

Denetimsiz öğrenmede (*unsupervised learning*) farklı olarak algoritmaya sadece girdi verisi sağlanır ve algoritmanın bu veriyi kendi kendine ayırması beklenir ⁹⁷. Bu

yöntem daha çok verideki gizli ilişkileri ve kalıpları ortaya çıkararak kümeleme yapmak için kullanılır.

Takviyeli öğrenme (*reinforcement learning*) de yine çıktılarının önceden tanımlanmadığı bir yöntemdir. Bunun yerine, model bir "ortam" içine yerleştirilir ve ortamla etkileşim kurması beklenir. Etkileşim sonucunda oluşan çıktılar modele “ödül” ya da “ceza” olarak geri verilir. Model, deneme-yanılma yoluyla kendini düzelterek ödülü maksimize, cezayı ise minimize etmeye çalışır⁹⁸. Bu yöntem daha çok robot kontrolü ve otomize sürüş gibi alanlarda kullanılır.

Bu 3 temel yöntem dışında, kendi kendine öğrenme, temel bileşen analizi, seyrek sözlük öğrenimi, anomali tespiti ve robot öğrenimi gibi yaklaşımlar da vardır. Bu yaklaşımların her birinin farklı güçlü ve zayıf yönleri olduğu için farklı alanlarda kullanılabilmektedir.

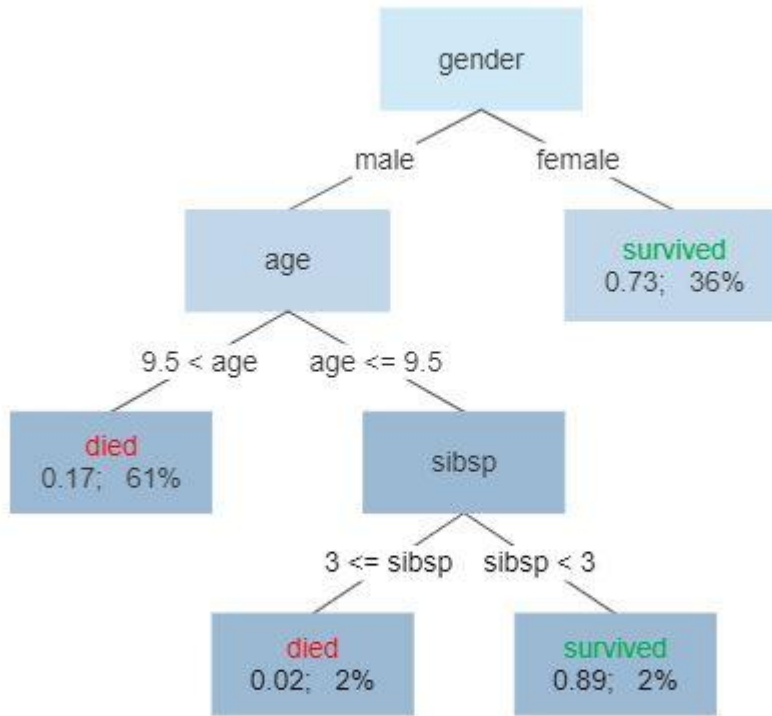
2.9.1. Karar Ağaçları (*Decision Trees* - DT) Algoritması

Karar ağaçları (*Decision Trees*, DT), denetimli öğrenme algoritmalarından biridir ve sınıflandırma problemlerinde kullanılır. Karar ağaçları, veri setindeki özellikler ile hedef değişken arasındaki ilişkileri modellemek için bir yapı sağlar. Bu ağaç yapısı, veriye dayalı kararlar almak ve yeni örneklerin sınıflandırılması veya hedef değerlerinin tahmin edilmesi için kullanılır.⁹⁹

Ağacın oluşturulması, kök düğümün seçimiyle başlar. Kök düğüm, veri setindeki en önemli özelliği veya değişkeni temsil eder. Bu seçim genellikle bilgi kazancı veya *gini impurity* gibi ölçütler kullanılarak yapılır. Bilgi kazancı, bir özelliğin kullanılmasıyla bilginin nasıl arttığını temsil ederken, *gini impurity* ise bir özelliğin kullanılmasıyla sınıflar arasındaki homojenliğin nasıl azaltıldığını ölçer.

Kök düğüm seçildikten sonra veri, seçilen özelliğe göre ayrılır ve 2 alt veri seti oluşturulur. Bu ayrıştırma işlemi, ağacın dallanması olarak adlandırılır. Oluşan her bir veri seti içinden, kök düğümünün seçimi ile aynı yöntem kullanılarak yeni bir düğüm seçilir. Seçilen her yeni düğüm sonrası dallanma ve düğüm seçme işlemi tekrar eder. Dallanma, bir sonlandırma kriteri (hiperparametre) karşılanana kadar, sonlanma kriteri olarak dal derinliği (kök düğümünden itibaren kaç kez dallandığı), dallanma öncesindeki veri setinin büyüklüğü, dallanma sonrası kalan veri setinin büyüklüğü, düğüm seçiminde kullanılacak özellik sayısı, en fazla yaprak düğüm sayısı gibi kriterler kullanılabilir. Bu kriterlerin seçimi kullanılan veri setine göre araştırmacı tarafından belirlenmelidir.

Survival of passengers on the Titanic



Şekil 4: Karar ağacı örneği

Görsel Wikipedia'dan alınmıştır¹⁰¹.

Oluşturulan ağaç yapısı zaman zaman aşırı karmaşık veya veriyle aşırı uyumlu olabilir. Bu durumda, ağaç yapısı düzeltilmelidir. Örneğin, aşırı dallanmış ağaçlar budanabilir veya düğümler birleştirilebilir. Budama işlemi, gereksiz dallanmaları ve

düğümlemleri kaldırarak ağacın karmaşıklığını azaltır ve genelleme yeteneğini artırır. Sonuç olarak, oluşturulan karar ağacı, yeni veri noktalarını sınıflandırmak veya hedef değerlerini tahmin etmek için kullanılabilir. Bir veri noktası, ağaçtaki karar düğümlerini takip ederek belirli bir sonuca ulaşır.

DT algoritması, veri setinin yapısını ve özelliklerin önemini anlama konusunda avantajlıdır. Ağaç yapısı, çeşitli durumları hızlı bir şekilde değerlendirebilir ve sonuçları kolayca yorumlanabilir hale getirir. Bu yorumlanabilirlik özelliği, DT algoritmasını tercih edilen bir seçenek haline getirir. Kullanıcılar, oluşturulan ağacı adım adım takip ederek karar sürecini anlayabilir ve elde edilen sonuçları açıkça yorumlayabilirler. Bu özellikler, DT'nin birçok uygulama alanında tercih edilmesini sağlar ¹⁰⁰.

2.9.2. Rastgele Orman (*Random Forest* - RF) Algoritması

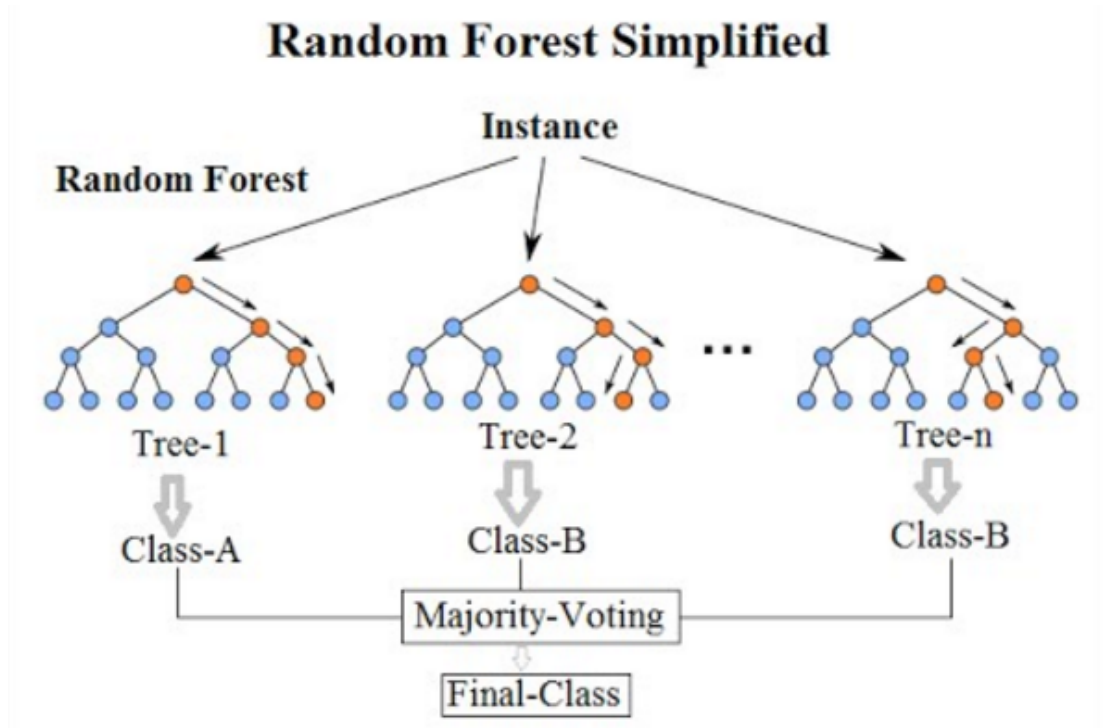
Rastgele Orman (*Random Forests* - RF), bir kombinasyon (*ensemble*) tabanlı sınıflama ve regresyon algoritmasıdır. Bu yöntemde, öncelikle veri seti içinden rastgele özellikler ve örnekler seçilerek bir alt küme oluşturulur. Bu alt küme, tek başına bir karar ağacı üreten DT algoritması ile değerlendirilir ve böylece bağımsız bir ağaç oluşturulur. Belirlenen sayıda ağaç oluşturana kadar bu işlemi tekrarlanır ¹⁰².

Tahmin etmek istediğimiz veri, oluşturulan bu ormandaki tüm ağaçlar tarafından değerlendirilir. Eğer sınıflandırma problemi için kullanıldıysa, her ağacın tahmini alınır ve en sık gözlenen sınıf tahmin modelin tahmini olarak kabul edilir. Eğer regresyon problemi için kullanıldıysa, her ağaç tarafından elde edilen tüm sonuçlar toplanır ve ardından bu sonuçların ortalaması alınarak modelin tahmini elde edilir.

RF, her ağacın kendi rastgele alt kümesiyle eğitilmesi nedeniyle bağımsızlığı sağlar ve bu, modelin daha iyi genelleme yapmasına ve aşırı uyuma riskinin

azaltılmasına katkıda bulunur ¹⁰³. Her ağaç, yalnızca veri setinin bir alt kümesini kullanarak öğrenir, böylece her ağaç kendine özgü bir bakış açısı ve farklı örüntüleri yakalar. Birçok ağacın birleşimiyle daha doğru ve kararlı bir sınıflandırma elde edilir.

RF algoritması, heterojen dağılıma sahip veri setleri veya eksik verilere sahip veri kümeleri gibi zorlu veri durumlarında da etkili bir şekilde kullanılabilir. Bu özellik, veri setindeki farklı yapı ve özelliklerin algoritmanın performansını artırabileceği anlamına gelir. Bu nedenle, veri setinin belirli özelliklerle homojen olmaması durumunda bile RF algoritması yüksek başarı oranları sağlayabilir.



Şekil 5: Rastgele orman algoritması örneği

Görsel Wikipedia'dan alınmıştır ¹⁰⁴.

Aynı zamanda, RF, gürültülü veriye ve aykırı değerlere daha az duyarlıdır. Çünkü her ağaç, rastgele alt kümelerle eğitildiğinden, bir ağacın yanlış bir aykırı değeri yakalaması diğer ağaçların daha sağlam ve doğru sonuçlar vermesini engellemez. Bu özellik, gerçek dünya problemlerinde RF'nin güvenilir ve etkili bir seçenek olarak tercih edilmesini sağlar.

Literatürde yapılan araştırmalar, RF algoritmasının çeşitli uygulama alanlarında başarılı sonuçlar verdiğini ve diğer sınıflandırma ve regresyon algoritmalarıyla karşılaştırıldığında önemli avantajlar sunduğunu göstermektedir ^{103,105}

2.9.3. Gradyan Artırmalı Karar Ağaçları (*Gradient Boosting Decision Trees* - GBDT) Algoritması

Gradyan Arttırma Karar Ağaçları (*Gradient Boosting Decision Trees* - GBDT), makine öğrenmesinde yaygın olarak kullanılan ve sınıflandırma ile regresyon problemlerinde etkili sonuçlar sağlayan bir başka kombinasyon öğrenme algoritmasıdır. Birden fazla karar ağacını bir araya getirerek güçlü bir tahmin modeli oluşturur ¹⁰⁶.

GBDT sınıflandırma problemleri için en sık gözlenen değer ve regresyon problemleri için ise ortalama değeri kullanan basit bir başlangıç modeliyle başlar. Bu başlangıç modeli, veri setinin örneklerine göre tahminler yapar ve elde edilen tahminler gerçek sonuçlarla karşılaştırır. Karşılaştırma sonucunda başlangıç hatası hesaplanır.

İkinci aşamada bu hataları tahmin etmek için DT algoritması çalıştırılır. Üretilen karar ağacı ile yeni bir tahmin yapılır ve yeniden hata hesaplanır. Yeni hesaplanan hata başlangıç hatasından çıkarılarak düzeltilmiş hata değerleri elde edilir. Daha sonra, bu düzeltilmiş hata değerleri temel alınarak tekrar DT algoritması çalıştırılır, yeni hata oranları elde edilir ve mevcut hata oranlarından çıkarılarak güncellenir.

Bu süreç, önceden belirlenen sayıda tekrar edilir ve bu tekrarlar sonucunda modelin eğitimi tamamlanır. Bu adımlarla, başlangıçta basit bir modelden

başlayarak, aşamalı olarak daha karmaşık ve doğru tahminler yapabilen bir model elde edilir.

GBDT, makine öğreniminde aşama aşama hatayı azaltarak bir hedefe doğru yaklaşan bir yöntemdir. Bu yöntem, ardışık sıralı hesaplamalara dayandığı için diğer kombinasyon öğrenme modellerine göre daha yüksek hesaplama gücüne ihtiyaç duymaktadır. Ancak, yapılandırılmış değişkenler içeren tablolaştırılmış veri setlerinde genellikle daha başarılı sonuçlar elde etmektedir.¹⁰⁷

2.9.4. Makine Öğrenimi Modellerini Değerlendirme Ölçütleri

Bir istatistik yönteminin ya da makine öğrenimi modelinin başarısının, matematiksel olarak değerlendirilebilmesi için bir çok ölçüt vardır. Ancak tüm bu ölçütlerin dayandığı üzere 4 temel değer bulunmaktadır¹⁰⁸.

- Gerçek Pozitif (True Positive - TP), gerçek durumu pozitif olan veri noktalarının doğru bir şekilde pozitif olarak sınıflandırıldığı durumları ifade eder.
- Gerçek Negatif (True Negative - TN), gerçek durumu negatif olan veri noktalarının doğru bir şekilde negatif olarak sınıflandırıldığı durumları ifade eder.
- Yanlış Pozitif (False Positive - FP), gerçek durumu negatif olan veri noktalarının yanlış bir şekilde pozitif olarak sınıflandırıldığı durumları ifade eder.
- Yanlış Negatif (False Negative - FN), gerçek durumu pozitif olan veri noktalarının yanlış bir şekilde negatif olarak sınıflandırıldığı durumları ifade eder.

Bu dört değer, testin veya modelin kullanıldığı alana bağlı olarak farklı oranlarda önem kazanır. Örneğin ilaç araştırmalarında, gerçekte etkili olmayan bir ilacın yanlışlıkla etkili gibi görünmesi, yani yanlış pozitif (FP) değerinin yüksek olması istenmeyen bir durumdur. Bu duruma Tip 1 hata da denir. Öte yandan kanser tarama testlerinde gerçekte kanseri olan hastaların yanlışlıkla sağlıklı gibi görünmesi, yani yanlış negatif (FN) değerinin yüksek olması istenmeyen bir durumdur. Buna da Tip 2 hata denir. Tip 1 ve Tip 2 hataların azaltılması için yukarıda bahsedilen dört temel ölçü arasındaki ilişkilere dayanarak yeni ölçütler geliştirilmiştir. Ancak, bu ölçütlerin tümü, testin amaçlandığı alana uygun şekilde seçilmeli ve yorumlanmalıdır.

Testlerin veya modellerin değerlendirilmesi için kullanılan diğer ölçütler aşağıda sıralanmıştır ¹⁰⁸⁻¹¹⁰:

- **Pozitif (Positive):** Gerçekten hasta olan tüm olguları temsil eder (TP + FN).
- **Negatif (Negative):** Gerçekten sağlıklı tüm olan olguları temsil eder (TN + FP).
- **Test Pozitif:** Testin pozitif dediği olguları temsil eder (TP + FP).
- **Test Negatif:** Testin negatif dediği olguları temsil eder (TN + FN).
- **Kesinlik (*Precision*):** Doğru pozitif tahminlerin tüm pozitif tahminlere oranını ifade eder. Kesinlik, FP değerlerin azaltılmak istendiği durumlarda kullanılır. (TP / (TP + FP)).
- **Duyarlılık / Geri çağırma / Gerçek Pozitiflik Oranı (*Sensitivity / Recall / True Positive Rate, TPR*):** Doğru pozitif tahminlerin pozitif olgulara oranını ifade eder. Duyarlılık, FN değerlerin azaltılmak istendiği durumlarda kullanılır (TP / (TP + FN)).
- **Yanlış Pozitiflik Oranı (*False Positive Rate, FPR*):** Yanlış pozitif tahminlerin negatif olgulara oranını ifade eder. FPR, FP değerlerin azaltılmak istendiği durumlarda kullanılır (FP / (TN + FP)). “1 - Özgüllük” olarak da ifade edilebilir.
- **Özgüllük / Gerçek Negatiflik Oranı (*Specificity, True Negative Rate, TNR*):** Doğru negatif tahminlerin negatif olgulara oranını ifade eder.

Özgüllük, FP değerlerin azaltılmak istendiği durumlarda kullanılır ($TN / (TN + FP)$). “1 - FPR” olarak da ifade edilebilir.

- **Doğruluk (Accuracy):** Doğru tahminlerin toplamının, toplam örneklem sayısına oranını ifade eder. Doğruluk, karşılaştırılan sınıfların dengeli dağıldığı durumlarda kullanılır. Ancak, dengesiz dağılım olan gruplarda yetersiz bilgi verebilir ($(TP + TN) / (TP + TN + FP + FN)$).
- **F1 skoru:** Duyarlılık ve Kesinlik değerlerinin geometrik ortalamasını temsil eder. Sınıflar arası dengesizliğin olduğu durumlarda kullanılır. ($2 * (Duyarlılık * Kesinlik) / (Duyarlılık + Keskinlik)$).
- **Jaccard indeksi:** İki kümenin kesişim kümesinin birleşim kümesine oranını verir. Kümelerin birbirine benzerliğini ifade eder. Sınıfların dengesiz dağıldığı durumlarda kullanılır. ($(|A \cap B|) / (|A \cup B|)$) ya da $TP / (TP + FP + FN)$).
- **Alıcı İşletim Karakteristiği (Receiver Operating Characteristic, ROC):** İkili sınıflandırma sistemlerinde duyarlılık ve özgüllük ilişkisinin gösterildiği eğriyi temsil eder. ROC eğrisi, x ekseninde FPR ve y ekseninde TPR temsil eder. Bir modelin ROC eğrisi, farklı sınırlayıcı değerler (*threshold*) kullanarak TPR ve FPR değerlerinin nasıl değiştiğini gösterir.
- **Eğri Altında Kalan Alan (Area Under Curve, AUC):** ROC eğrisinin altında kalan alanı ifade eder. AUC, bir sınıflandırma modelinin performansını tek bir sayısal değerle ölçer. AUC değeri 0 ile 1 arasında değişir ve daha yüksek bir AUC değeri daha iyi bir performans gösterir. AUC değeri 0.5'e yakınsa, modelin performansı rastgele sınıflandırma seviyesine yakındır. AUC değeri 1'e yaklaştıkça, modelin performansı artar ve mükemmel bir sınıflandırma yeteneği gösterir. AUC, sınıfların dengesiz dağıldığı durumlarda daha güvenilir bir performans ölçütü olarak kullanılır. Modelin sınıflamada ne kadar başarılı olduğunu gösterir.

Kullanılacak ölçüt, veri setinin dağılımına ve öncelenen çıktıya göre araştırmacı tarafından seçilmelidir. Bu seçim, testin veya modelin performansını en iyi şekilde değerlendirebilmek için önemlidir. Farklı ölçütlerin kullanılması, testin veya modelin performansının farklı açılardan incelenmesine yardımcı olabilir.

Araştırmacılar genellikle farklı ölçütleri kullanarak testleri veya modelleri tekrar tekrar çalıştırır ve performanslarını değerlendirir. Bu süreç, algoritmaların farklı ayarlarını (hiperparametre) deneyerek en iyi sonuçları elde etmeyi amaçlar. Ölçütler arasındaki karşılaştırma, testin veya modelin hedeflenen amaca en uygun şekilde hizmet etmesini sağlamak için önemlidir. Bu süreç, daha güvenilir sonuçlar elde etmeyi ve test veya modelin etkinliğini artırmayı sağlar.

2.9.5. Veri Dağılımı ve Eksik Veriler

Makine öğrenimi alanında güvenilir sonuçlar elde etmek için, karşılaştırılacak veri setlerinin benzer büyüklükte olması önemlidir. Bir tarafın diğerinden belirgin şekilde daha küçük olması durumunda, elde edilen sonuçların tahmin gücü azalmaktadır. Veri dengesizliğini azaltmak için ilk adım, az miktarda veri içeren sete yeni verilerin eklenmesidir. Ancak, bu yöntem her zaman uygulanabilir değildir. Örneğin, nadir bir hastalığa sahip hastaların sayısı zaten sınırlı olduğundan, daha fazla hasta bulmak mümkün olmayabilir.

Bu gibi durumlarda, veri setlerini dengelemek için iki yöntem kullanılabilir: yetersiz örnekleme (*undersampling*) veya aşırı örnekleme (*oversampling*). Yetersiz örneklemede, büyük olan tarafın azaltılmasıyla veri seti dengelenirken, aşırı örneklemede küçük olan tarafın örnekleri çoğaltılarak veri seti genişletilir. Yapılan çalışmalar, aşırı örnekleme yöntemlerinin, makine öğreniminde daha başarılı sonuçlar sağladığını göstermektedir ¹¹¹.

Aşırı örnekleme için birden fazla yöntem bulunmaktadır, ve bu yöntemlerden en basiti rastgele aşırı örneklemedir ¹¹². Bu yöntemde, küçük veri setinden rastgele örnekler seçilerek kopyalanır, böylece veri sayısı dengelenir. Veri setinin

büyütülmesiyle birlikte bilgi kaybı oluşmaz, ancak modelin kopyalanan verilere aşırı uyum sağlama eğilimi vardır.

Azınlığın Yapay Aşırı Örneklemme Tekniği (*Synthetic Minority Over-sampling Technique*, SMOTE), rastgele kopyalamanın zayıf yönlerini gidermek amacıyla geliştirilmiştir ¹¹². Bu teknikte, veri setinden rastgele bir örnek seçilir ve bu noktaya yakın bir nokta belirlenir. İki nokta arasındaki özellik vektörü hesaplanır. Bu özellik vektörü, 0 ile 1 arasında bir ağırlıkla çarpılarak yeni bir özellik vektörü oluşturulur. Bu yeni özellik vektörüne sahip olan veri noktası, veri setine eklenir.

SMOTE yöntemiyle üretilen yapay veriler, mevcut verilere göre çok az bilgi farklılığı içerir. Ayrıca, yapay verilerin model tarafından aşırı uyum sorununa yol açması daha az olasıdır. Üretilen noktalar, var olan noktaların komşularına göre belirlendiği için, verilerin seyrek dağıldığı bölgelerde yapay veri üretmek zor olabilir. Bununla birlikte, yapay olarak üretilen verilerin her zaman gerçek durumla tam olarak uyuşmadığı da unutulmamalıdır. SMOTE, literatürde en yaygın kullanılan aşırı örneklemme yöntemidir ve farklı analizler için birçok farklı versiyonu üretilmiştir

¹¹³.

Veri analizinde, eksik verilerin yönetimi de bir başka önemli konudur. Eksik veriler, sonuçların güvenilirliğini etkileyebilir ve bazı algoritmalar eksik verilerle başa çıkmada kısıtlamalar yaşayabilir. Örneğin, k Yakın Komşular algoritması eksik verilerle çalışabilir ve eksik verilere dayalı tahminler yapabilir. Ancak, RF ve GB algoritmaları eksik verilerle çalışamaz ve eksik verilerin yönetimi için farklı yaklaşımlar gerektirebilir ¹¹⁴.

İdeal durumda, tüm verilerin eksiksiz bir şekilde toplanması tercih edilir. Ancak, pratikte bu ideal durum her zaman gerçekleştirilemeyebilir. Bu gibi durumlarda, eksik verilerin veri setinden çıkarılması, güvenilir sonuçlar elde etmek için tercih edilen bir yaklaşımdır. Özellikle nadir hastalık araştırmaları gibi durumlarda, eksik verilerin çıkarılması, zaten sınırlı olan örneklem sayısını daha da azaltabilir. Bu

nedenle, eksik verilerin yerine tahmin veya hesaplama yöntemleriyle yeni veriler atanması (*impute*) seçilebilecek bir yöntemdir.

Eğer eksik veri sayısal bir değer ise, eksik değerlerin doldurulması için veri dağılımına göre ortalama veya ortanca değerler kullanılabilir. Eğer veri seti normal dağılım gösteriyorsa, ortalama değer kullanımı daha uygundur, ancak normal dağılım göstermiyorsa ortanca değer kullanılması tercih edilebilir ¹¹⁵.

Kategorik değişkenlerde ise, eksik verilerin doldurulması için farklı bir yaklaşım izlenir. Eğer kategoriler birbiriyle eşit önemde ise (örneğin kan grubu), en sık gözlenen değer kullanılabilir. Ancak, kategoriler arasında bir hiyerarşi bulunuyorsa (örneğin Likert ölçeği), "normal" olarak kabul edilen bir değer kullanılabilir ¹¹⁵.

Makine öğreniminde, model ve ölçüt seçiminde de olduğu gibi, dengesiz grupların ve eksik verilerin değerlendirilmesi konusunda kesin bir yöntem yoktur. Kullanılacak yöntem veri setinin boyutu, gruplar arasındaki fark büyüklüğü, değişkenlerin özelliği, eksik değişkenlerin sayısı, veri dağılımı ve modelin kullanım amacı göz önünde bulundurularak seçilmelidir.

2.10. Genetik

Genetik, görece yeni bir bilim sahası olmasına rağmen, insanlar yüzyıllar öncesinde bile bazı özelliklerin soy boyunca devam ettiğini gözlemlemiştir. Bu aktarımın nasıl olduğuna dair yazılı ilk kaynaklar Hipokrat ve Aristoteles'e aittir ¹¹⁶. Modern anlamda genetik kalıtımın temel yasaları, 1866 yılında Gregor Mendel tarafından açıklanmıştır ¹¹⁷.

1900'lü yılların başında kromozomlar keşfedilmiş, 1914 yılında "genotip" ve "fenotip" kavramları literatüre eklenmiştir ¹¹⁸. DNA ve nükleik asitlerin yapısı ise

1954 yılında ise James Watson ve Francis Crick tarafından keşfetmiştir ¹¹⁹. Bu keşifle birlikte moleküler genetik çağı başlamıştır.

Çeşitli bakteri ve virüslerin genomunu başarıyla dizileyen bilim insanları, 1990 yılında, İnsan Genom Projesi'ni başlatmışlardır. İnsan genomunun %92'si 2003 yılında, %99,7'si 2021 yılında ve tamamı ise 2022 yılında başarıyla dizilenmiştir ¹²⁰. Bu çalışmalar, gen dizileme teknolojisinin hızla gelişmesine ve maliyetinin düşmesine yol açmıştır. Sonuç olarak, genetik tanı yöntemlerine erişim artmıştır. Aynı zamanda, birçok yeni moleküler yolak ve hastalık keşfedilmesi de mümkün olmuştur. Günümüzde genetik testler, yenidoğan taraması, implantasyon öncesi tarama, taşıyıcı taraması, hastalık teşhisi, farmakogenetik, adli inceleme ve genetik soyağacı çıkartma gibi tıbbi ve tıp dışı birçok alanda kullanılmaktadır.

Genetik bozukluklar, tek nükleotid değişimi (*Single Nucleotide Polymorphism*, SNP) gibi basit değişikliklerden kromozomun kısmi ya da tamamen silinmesi, kopması, çoğalması hatta bütün bir kromozomun çiftleşmesi gibi köklü değişikliklere kadar çeşitlilik gösterebilir. Bu farklı mutasyon çeşitlerinin tespiti için Polimeraz Zincir Reaksiyonu (*Polymerase Chain Reaction*, PCR), genom dizileme, sitogenetik ve mikro dizileme (*microarray*) gibi birbirinden farklı genetik testler kullanılmaktadır. Her bir testin güçlü ve zayıf yönleri bulunmaktadır. Amaca uygun yöntemi belirlemek, uzmanlık gerektiren bir karardır. Her yöntem, analiz kapasitesi, hassasiyet, örnek gereksinimleri ve maliyet gibi faktörler açısından değerlendirilmelidir. Bu şekilde, amaç ve gerekliliklere en uygun yöntem seçilerek güvenilir sonuçlar elde edilebilir.

PCR, belirli bir gen bölgesinin çoğaltılması işlemidir. Bu yöntem, az miktarda genetik materyalin artırılmasıyla üzerinde daha detaylı çalışmaların yapılabilmesine imkan tanır. Bunun yanında bilinen bir genetik mutasyonun tespiti için de kullanılabilir ^{121,122}.

DNA dizileme, seçilen genetik bölgedeki nükleotidlerin sıralamasını belirlemek için kullanılan bir yöntemdir. Bu yöntemde, nükleotitler DNA'da bulundukları sıra

ile tespit edilebilir. Tespit edilen deęişiklikler literatür ile karşılaştırılabilir, hastalık ile ilişkili mutasyonlar incelenebilir.

Keşfedilen ilk dizileme yöntemi olan Sanger dizileme, hala kullanılmakta ve altın standart olarak kabul edilmektedir. Bu yöntemde, floresan işaretli nükleotitler, hedeflenen DNA'nın dizilenmek istendięi ortama eklenir. İşaretli nükleotitler, DNA polimeraz enzimi aracılığıyla replikasyon sırasında normal nükleotitlerin yerini alır ve replikasyonu durdurur. Bu işlem sonucunda, farklı uzunluklarda DNA parçaları elde edilir. Daha sonra, elektroforez yöntemiyle bu DNA parçaları büyüklüklerine göre ayrıştırılır ve floresan ışımaya maruz bırakılarak içerdikleri nükleotit dizileri sırasıyla okunabilir ¹²³. HGP kapsamındaki dizilemeler bu yöntemle yapılmıştır ¹²⁴. Sanger dizileme, güvenilir bir yöntem olmakla birlikte, yavaş çalışması ve yoğun emek gerektirmesi nedeniyle sınırlamalara sahiptir. Ayrıca, 300-1000 baz çifti uzunluęundaki DNA parçalarıyla çalışmayı gerektirmesi de yöntemin zayıf yönlerinden biridir.

Yeni Nesil Dizileme (*Next-generation sequencing*, NGS), Sanger dizilemesinden daha hızlı çalışan birkaç teknolojinin tamamına verilen bir isimdir ¹²⁵. İkinci nesil dizileme olarak da adlandırılmaktadır. Bu yöntemlerin ortak özellięi, aynı anda birden fazla DNA parçasının dizilenmesini mümkün kılmasıdır ¹²⁶. HGP kapsamında Sanger sekanslama ile tek bir kişinin genom dizilemesi 10 yılda fazla sürmüştür. NGS yöntemleri ile tüm genom dizilemesi yalnızca 1 gün sürmektedir ¹²⁷. Bu hızlı dizileme süresi, NGS'nin genetik araştırmalarda ve klinik uygulamalarda geniş bir kullanım alanı bulmasını sağlamıştır (Goodwin et al., 2016).

NGS yöntemi kullanılarak Tüm Genom Dizilemesi (*Whole Genome Sequencing*, WGS) yapılabilmesinin yanı sıra, sadece protein kodlayan egzomları içeren Tüm Egzom Dizilemesi (*Whole Exome Sequencing*, WES) de gerçekleştirilebilir ¹²⁸. WES yöntemi, genomun sadece %1'lik kesimini taradıęı için WGS'ye kıyasla daha ekonomik ve hızlı bir yöntemdir ¹²⁹. NGS kullanılarak bilinen mutasyonlar taranabilirken aynı zamanda yeni mutasyonlar da keşfedilebilir ¹³⁰. NGS ile elde edilen tüm mutasyonlar, hastalık sebebi olmak zorunda değildir. Patojenik

mutasyonların tespiti için klinisyenler tarafından, OMIM gibi veritabanları kullanılarak, hastanın fenotipik özelliklerinin değerlendirilmesi gerekmektedir.

Gelişen teknolojiye rağmen, WES ve WGS yöntemleri hala yavaş, maliyetli ve zorlayıcıdır. Dizileme için özel laboratuvarlar ve deneyimli personel gereklidir. Ayrıca, dizileme sonuçlarının yorumlanması da önemlidir. Bu yöntemler, yeni mutasyonların keşfi konusunda başarılı olsa da, bilinen mutasyonların taranması için yeterince pratik değildir.

NGS'in en güçlü yönü, paralel dizileme yeteneğidir. Bu özellik, birden fazla genetik bölgenin aynı anda taranabilmesini sağlar. NGS yöntemi, hastalıkla ilişkili literatürde iyi tanımlanmış genlerin ve mutasyonların taranması için kullanılabilir. Bu amaçla, taranan genlerin bulunduğu çalışmalara "panel dizileme" denir. Panel dizileme, hızlı, düşük maliyetli ve pratik bir yöntemdir. Sonuçları daha kolay yorumlanabilir ve klinik uygulamalarda daha uygun bir seçenektir ¹³¹.

Ancak, panel dizilemenin en zayıf yönü, sadece sınırlı sayıda mutasyonu taraması nedeniyle yeni tanımlanan ve henüz panele dahil edilmemiş mutasyonları gözden kaçırabilmesidir. Bununla birlikte, panel dizileme hızla gelişmektedir ve yeni mutasyonların sistematik olarak eklenmesi ve güncellenmesi ile bu kısıtlama azaltılmaktadır. Panel dizileme yöntemi sürekli olarak güncellenen genetik veritabanları ve literatür bilgileriyle entegre edilerek, yeni mutasyonları tespit etme yeteneği artırılmaktadır.

Sitogenetik, kromozomların morfolojisinin incelenmesidir. Kromozomların tamamının incelendiği bir teknik olan karyotipleme, bu amaçla kullanılan bir yöntemdir ¹³². Genetik cinsiyet tespiti, monozomiler, trizomiler ve bazı translokasyonlar gibi durumlar bu yöntemle tespit edilebilir. Yerinde Floresan Melezleme (*Fluorescence in situ Hybridization*, FISH) yöntemi ise kromozomun belirli bir bölgesine seçici olarak bağlanabilen boyaların kullanıldığı bir yöntemdir ¹³³. Bu yöntem, yer değiştirme, silinme ve çoğalma gibi mutasyonların tespitini yapabilir. FISH, hücrelerdeki belirli genlerin konumlarını ve sayılarını belirlemek,

kromozomal anormallikleri teşhis etmek ve bazı genetik hastalıkların tanısını koymak için yaygın olarak kullanılan bir yöntemdir.

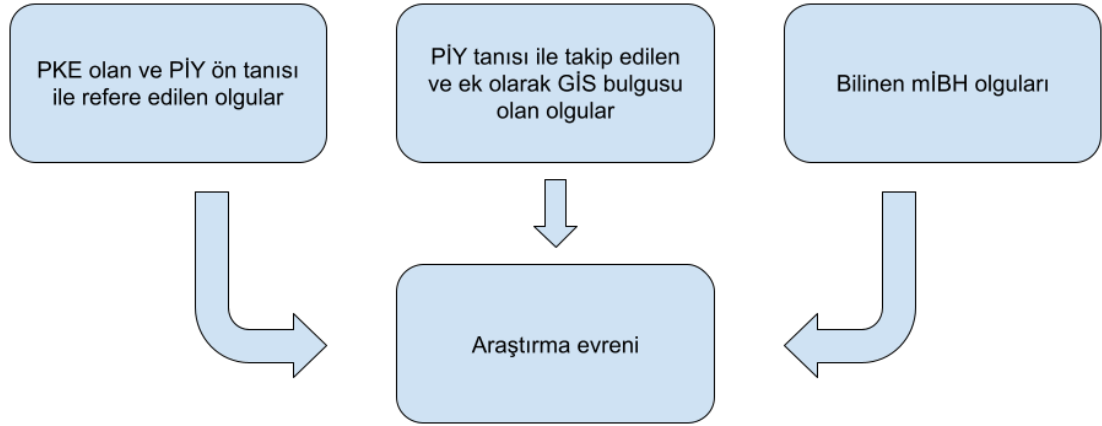
Mikrodizileme, kromozomun tamamını incelerken FISH veya karyotip ile tespit edilemeyecek kadar küçük değişiklikleri tespit etme yeteneğine sahiptir. Bu teknikte, yapay olarak üretilen binlerce kısa, tek zincirli DNA dizisini içeren çipler kullanılır. Bu diziler, insan genomunda bulunan genler veya gen varyasyonları olabilir. Taranması amaçlanan genetik materyal, floresan boyalarla işlem gördükten sonra bu çiplere eklenir. Eğer çip üzerindeki genetik dizi, taranan materyalle uyumlu ise boyalı genetik materyal çipe bağlanır¹³⁴. Mikrodizileme ile yüzlerce mutasyon aynı anda taranabilir. Bir genetik bölgenin varlığı ya da yokluğu tespit edilebilir. Mikrodizileme ile DNA'nın tam dizilemesi yapılamaz, ama mutasyon olan bölgenin varlığını gösterebilir. Bu nedenle de karyotipleme ile birlikte tercih edilmesi gereken ilk basamak testler içinde yer alır.

3. GEREÇ VE YÖNTEM

3.1. Araştırma Evreni

Bu tez, Marmara Üniversitesi Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Çocuk Alerji ve İmmünoloji Bilim Dalı tarafından yürütülen dört adet prospektif çalışmaya dahil olan hastalar temel alınarak gerçekleştirildi. Toplamda 520 olgu değerlendirildi. Çalışmada araştırma evrenine, PKE tanısı olan ve immün yetmezlik ön tanısı ile kliniğimize yönlendirilen hastalar; PİY tanısı ile takip edilen ve ek olarak GİS bulguları görülen hastalar ve de bilinen mİBH olan olgular dahil edildi.

Olguların verileri, hastane otomasyon sistemi, bölümdeki hasta dosyaları ve önceki çalışmalar kapsamında toplanan verilerden erişildi. Detaylı klinik bilgisine ulaşamayan hastalar çalışma dışı bırakıldı. Bunun yanında genetik incelemesi (panel, WES ya da WGS) yapılmayan hastalar da araştırma evreninden çıkarıldı.



Şekil 6: Araştırmaya dahil edilen hastaların seçildiği kaynaklar

3.2. İzinler ve Onamlar

Bu tez, mevcut hasta dosyaları ve verileri üzerinden yürütülmüştür. Hastalara ek işlem yapılmamış, herhangi bir ek tetkik istenmemiştir. Çalışma, Marmara Üniversitesi Tıp Fakültesi Klinik Araştırmalar Etik Kurulu tarafından onaylanmıştır. (Protokol no: 09.2022.1080) (EK-1)

3.3. Veri Toplama

Tez kapsamında, hastaların bulguları, semptomları ve laboratuvar sonuçları, HPO Nisan 2023 versiyonunda mevcut olan terimler kullanılarak kodlandı. Eğer kullanılan versiyonda bulunmayan terimler mevcutsa, bu terimler olduğu şekliyle eklendi.

Araştırma evrenine dahil olan hastaların demografik bilgiler, hastaların başlangıç muayenesinden başlayarak tanı konulana ve de varsa, tedavi başlayana kadar olan bulgu, semptom ve laboratuvar bulguları çalışmaya dahil edildi. Fizik muayene bulguları, laboratuvar verileri (albumin, total protein, nutrisyonel parametreler, transaminazlar, böbrek fonksiyon testleri, elektrolit bozuklukları vb) ve varsa, batin görüntülemelerine ilişkin bulgular, literatürde önerilen yöntemle ⁸⁷, HPO terimleri kullanılarak veritabanına eklendi.

Hastaların tercihen ilk geliş muayenelerindeki, yoksa, erişilebilen en eski boy ve kilo bilgileri, ölçüm tarihi ile birlikte, sayısal olarak girildi. Daha sonra z-skorları `rcpchgrowth-python` ¹³⁵ kütüphanesi ile hesaplanılıp -2 ile +2 aralığı dışında kalanlar HPO terimleri veritabanına eklendi.

Hastaların tam kan sayımı ve Ig değerleri, örneklem tarihleri ile birlikte sayısal olarak girildi. Değerler, yaşa göre referanslar ^{136,137} baz alınarak, Python dilinde yazılan bir betikle hesaplandı. Normal dışı parametrelere ait HPO terimleri de verilere eklendi.

Hastaların lenfosit alt grupları da yine yaşa göre referans değerler ¹³⁸ baz alınarak hesaplandı. Sonuçlar “Düşük”, “Normal” ve “Yüksek” olarak kodlanarak veri setine eklendi. Laboratuvarımızda bakılan B ve T lenfosit alt gruplarının tamamı mevcut HPO tanımları arasında yer almadığı için HPO terimleri kullanılamadı.

3.4. Analiz

Verilerin analizi ve görselleştirmesi için IBM SPSS 29, Python 3.10 ¹³⁹, Pandas 2.0.1 ¹⁴⁰, scikit Learn 1.3.0 ¹⁴¹, imbalanced-learn 0.10.1 ¹⁴² ve Plotly 5.14.1 ¹⁴³ programları ve kütüphaneleri kullanıldı. Değişkenlerin normallik dağılımlarını değerlendirmek için Shapiro-Wilk (S-W) testi kullanıldı. Süreğen değişkenler normal dağılım gösteriyorsa ANOVA testi ile karşılaştırıldı. İkili grupların analizi için ise Tukey HSD testi kullanıldı. Normal dağılım göstermeyen süreğen değişkenler ve tüm kategorik değişkenler için ise Kruskal-Wallis ve *post-hoc* testleri kullanıldı. Sonuçların değerlendirilmesi için Bonferroni düzeltmesi yapılmış p değerleri kullanıldı.

Hastaların klinik ve laboratuvar özelliklerine göre sınıflandırma işlemi için makine öğrenimi yöntemleri kullanıldı. Çalışılan veri setinin özgün olması nedeniyle literatürde sık kullanılan, doğrusal olmayan aktivasyon fonksiyonlarına sahip olan ve farklı senaryolarda iyi performans gösterdiği bilinen RF ve GBDT algoritmaları tercih edildi.

Hasta sayısının az olması nedeniyle algoritmalar, hem veri setinin orjinal haliyle, hem de SMOTE yöntemi ile dengelenmiş haliyle çalıştırıldı. Başarılı ayırım yapabilen modellere ulaşabilmek için seçilen algoritmaların farklı hiperparametre kombinasyonlarda seçildi. Modeller, aynı veri seti içinden rastgele seçilen, birbirinden farklı 3 eğitim ve deneme parçası ile eğitildi.

Model başarısını değerlendirmek için f1 skoru, ROC - AUC ve Jaccard indeksi kullanıldı. Seçilen ölçütünün skoru 0,75'den büyük olan modeller, çıktıları incelenmek üzere, parametleri ile birlikte kaydedildi.

Tez içinde kullanılan grafiklerin etkileşimli versiyonlarına aşağıdaki QR kodu ya da <https://birkanazer.github.io/residency-thesis/> üzerinden erişebilirsiniz.



Şekil 7: İnteraktif grafikler için QR kodu

4. BULGULAR

Araştırma evreni, kliniğimizde gerçekleştirilen prospektif çalışmaların hastalarının verilerini içermektedir. Toplamda 520 hasta üzerinde yapılan taramalar sonucunda elde edilen veriler değerlendirilmiştir. Bu değerlemeye göre hastaların %66,15'inde (n = 344) klinik olarak anlamlı genetik mutasyonlar tespit edilmiştir. Bu bulgular ışığında, en sık gözlenen üç mutasyonun sırasıyla *LRBA* (%10,58), *CD55* (%8,08) ve *CTLA4* (%5,38) olduğu belirlenmiştir. Elde edilen tüm mutasyon verileri Tablo 3'te ayrıntılı olarak sunulmuştur.

Tablo 3: Tüm kohortlarda genetik etyoloji dağılımı

Gen	Frekans	Gen	Frekans	Gen	Frekans	Gen	Frekans
<i>LRBA</i>	%10,58 (55)	<i>CCBE1</i>	%0,58 (3)	<i>PMM2</i>	%0,19 (1)	<i>TYMP</i>	%0,19 (1)
<i>CD55</i>	%8,08 (42)	<i>IRHOM2</i>	%0,58 (3)	<i>PIEZO1</i>	%0,19 (1)	<i>TNFRSF13B</i>	%0,19 (1)
<i>CTLA4</i>	%5,38 (28)	<i>PIK3CD</i>	%0,58 (3)	<i>NCF4</i>	%0,19 (1)	<i>IGHM</i>	%0,19 (1)
<i>BTK</i>	%2,69 (14)	<i>RAG2</i>	%0,58 (3)	<i>MPI</i>	%0,19 (1)	<i>SP110</i>	%0,19 (1)
<i>RAG1</i>	%2,5 (13)	<i>STAT1 GOF</i>	%0,58 (3)	<i>FAT4</i>	%0,19 (1)	<i>ALG6</i>	%0,19 (1)
<i>CYBB</i>	%2,5 (13)	<i>IL2RG</i>	%0,58 (3)	<i>EPCAM</i>	%0,19 (1)	<i>ST6GALNAC1</i>	%0,19 (1)
<i>NCF1</i>	%2,31 (12)	<i>ADAMTS3</i>	%0,38 (2)	<i>JAK3</i>	%0,19 (1)	<i>IL37</i>	%0,19 (1)
<i>DNMT3B</i>	%1,92 (10)	<i>FERMT1</i>	%0,38 (2)	<i>TRAF3IP2</i>	%0,19 (1)	<i>POMP</i>	%0,19 (1)
<i>STAT3 LOF</i>	%1,73 (9)	<i>ARPC1B</i>	%0,38 (2)	<i>IL21R</i>	%0,19 (1)	<i>PRKCD</i>	%0,19 (1)
<i>CARMIL2</i>	%1,54 (8)	<i>PIK3C2A</i>	%0,38 (2)	<i>RFX5</i>	%0,19 (1)	<i>IL2RB</i>	%0,19 (1)
<i>FOXP3</i>	%1,54 (8)	<i>ZBTB24</i>	%0,38 (2)	<i>PEPD</i>	%0,19 (1)	<i>STAT3 GOF</i>	%0,19 (1)
<i>CD40LG</i>	%1,35 (7)	<i>ZAP70</i>	%0,38 (2)	<i>AICDA</i>	%0,19 (1)	<i>IL10RB</i>	%0,19 (1)
<i>ALB</i>	%1,15 (6)	<i>MALT1</i>	%0,38 (2)	<i>GIMAP5</i>	%0,19 (1)	<i>IL10RA</i>	%0,19 (1)
<i>CYBA</i>	%1,15 (6)	<i>WAS</i>	%0,38 (2)	<i>ADA2</i>	%0,19 (1)	<i>STAT1</i>	%0,19 (1)
<i>ADA</i>	%1,15 (6)	<i>TTC7A</i>	%0,19 (1)	<i>CCNO</i>	%0,19 (1)	<i>IL2RA</i>	%0,19 (1)
<i>G6PC3</i>	%1,15 (6)	<i>FCHO1</i>	%0,19 (1)	<i>DYSF</i>	%0,19 (1)	<i>SKIV2L</i>	%0,19 (1)
<i>DGAT1</i>	%0,96 (5)	<i>YARS</i>	%0,19 (1)	<i>POLE2</i>	%0,19 (1)	Mutasyon tespit edilemedi	%33,85 (176)
<i>MEFV</i>	%0,96 (5)	<i>DPAGT1</i>	%0,19 (1)	<i>ICOS</i>	%0,19 (1)		
<i>NCF2</i>	%0,96 (5)	<i>TTC37</i>	%0,19 (1)	<i>CD19</i>	%0,19 (1)		
<i>PIK3R1</i>	%0,77 (4)	<i>SLCO2A1</i>	%0,19 (1)	<i>STXBP2</i>	%0,19 (1)		
<i>DCLRE1C</i>	%0,77 (4)	<i>XIAP</i>	%0,19 (1)	<i>PGM3</i>	%0,19 (1)		

Tablo 4: Genlerin DBK ve mİBH sınıflamalarına göre dağılımı

I. mİBH ve DBK	<i>ADA</i>	<i>CYBB</i>	<i>IL2RB</i>	<i>PIK3CD</i>	<i>TGFB1</i>
	<i>ADA2</i>	<i>CYBC1</i>	<i>IL2RG</i>	<i>PIK3R1 *</i>	<i>TGFBRI</i>
	<i>ALPI</i>	<i>DCLRE1C</i>	<i>IRF2BP2</i>	<i>POLA1</i>	<i>TGFBRI2</i>
	<i>ARPC1B</i>	<i>DKC1</i>	<i>ITCH</i>	<i>PRKCD</i>	<i>TNFAIP3</i>
	<i>BACH2</i>	<i>DOCK2</i>	<i>ITGB2</i>	<i>RAG1</i>	<i>TRIM22</i>
	<i>BCL10</i>	<i>ELF4</i>	<i>LIG4</i>	<i>RAG2</i>	<i>TRNT1</i>
	<i>BTK</i>	<i>FERMT1</i>	<i>LRBA</i>	<i>RBCK1</i>	<i>TTC37</i>
	<i>CARD11</i>	<i>FOXP3</i>	<i>MALT1</i>	<i>RELA</i>	<i>TTC7A</i>
	<i>CARMIL2</i>	<i>G6PC3</i>	<i>MASP2</i>	<i>RIPK1</i>	<i>WAS</i>
	<i>CASP8</i>	<i>ICOS</i>	<i>MVK</i>	<i>RTEL1</i>	<i>WIPF1</i>
	<i>CD3G</i>	<i>IKBKG</i>	<i>NCF1</i>	<i>SAMD9</i>	<i>XIAP</i>
	<i>CD40LG</i>	<i>IL10</i>	<i>NCF2</i>	<i>SKIV2L</i>	<i>ZAP70</i>
	<i>CD55 *</i>	<i>IL10RA</i>	<i>NCF4</i>	<i>STAT1</i>	
	<i>CDC42</i>	<i>IL10RB</i>	<i>NFAT5</i>	<i>STAT3 LOF</i>	
	<i>CTLA4</i>	<i>IL21</i>	<i>NFKBIA</i>	<i>STIM1</i>	
	<i>CYBA</i>	<i>IL2RA</i>	<i>NLRC4</i>	<i>SYK</i>	
II. mİBH, DBK değil	<i>ANKZF1</i>	<i>HPS1</i>	<i>IL37 *</i>	<i>PTPN2</i>	<i>SLC9A3</i>
	<i>CARD8</i>	<i>HPS3</i>	<i>LACC1</i>	<i>SCGN</i>	<i>SLCO2A1</i>
	<i>COL7A1</i>	<i>HPS4</i>	<i>NOP10</i>	<i>SIRT1</i>	<i>STXBP3</i>
	<i>FMNL2</i>	<i>HPSA1L</i>	<i>NPC1</i>	<i>SLC26A3</i>	<i>TYMP</i>
	<i>GUCY2C</i>	<i>IKZF1</i>	<i>PI4KA</i>	<i>SLC37A4</i>	<i>WNT2B</i>
III. mİBH değil, DBK	<i>AICDA</i>	<i>FAT4</i>	<i>JAK3</i>	<i>POLE2</i>	<i>STAT3 GOF</i>
	<i>CCBE1</i>	<i>FCHO1</i>	<i>MEFV</i>	<i>RFX5</i>	<i>TNFRSF13B</i>
	<i>CD19</i>	<i>IGHM</i>	<i>PEPD</i>	<i>SP110</i>	<i>TRAF3IP2</i>
	<i>DNMT3B</i>	<i>IL21R</i>	<i>PGM3</i>	<i>STAT1 GOF</i>	<i>ZBTB24</i>
IV. Sınıflandırılmamış Mutasyon	<i>ADAMTS3</i>	<i>DGAT1</i>	<i>GIMAP5</i>	<i>PIK3C2A</i>	<i>YARS</i>
	<i>ALB</i>	<i>DPAGT1</i>	<i>IRHOM2 *</i>	<i>PMM2</i>	
	<i>ALG6</i>	<i>DYSF</i>	<i>MPI</i>	<i>POMP</i>	
	<i>CCNO</i>	<i>EPCAM</i>	<i>PIEZO1</i>	<i>ST6GALNAC1 *</i>	

Taranan kohortlardaki genetik mutasyonları farklı sınıflama sistemlerine göre gösterimi

I: Güncel literatüre göre mİBH¹ etkeni olarak tanımlanan aynı zamanda da DBK⁴ olarak sınıflanan genler.

II: Güncel literatüre göre mİBH etkeni olarak tanımlanan ancak DBK sınıflamasına dahil edilmemiş olan genler

III: Tez kapsamında taranan kohortlarda saptanan DBK sınıflamasına dahil olan ancak mİBH etkeni olarak tanımlanmayan genler

IV: Güncel literatüre göre mİBH ve DBK etkenleri arasında yer almayan genler.

*: İlk kez kliniğimizdeki çalışmalarda tanımlanan mutasyonlar.

Kalın: Kliniğimizde yürütülen kohort çalışmalarında takip edilen hastalar.

Kohortlarımızda bulunan genlerden bazıları literatürde mİBH etkeni olarak, bazıları da DBK etkeni olarak sınıflanmaktadır. Bazı genler her iki listede de yer almaktayken bazı genler her iki listenin de dışında bulunabilmektedir. Taranan kohortlarda mevcut olan mutasyonların tamamının farklı sınıflama sistemlerinde dağılımı Tablo 4'te özetlenmiştir.

Çalışmamızın belirlediği dahil etme ve hariç tutma kriterlerine uyan 200 hasta, araştırma evrenine dahil edilmiştir. Bu hasta grubunun %36'sı (n = 72) kız, %64'ü ise (n = 128) erkek cinsiyetine sahiptir. Hastaların %61,1'inde (n = 118) akraba evliliği durumu tespit edilmiştir. Hastaların demografik özellikleri, Tablo 5'te ayrıntılı olarak sunulmuştur.

Tablo 5: Hastaların demografik dağılımları

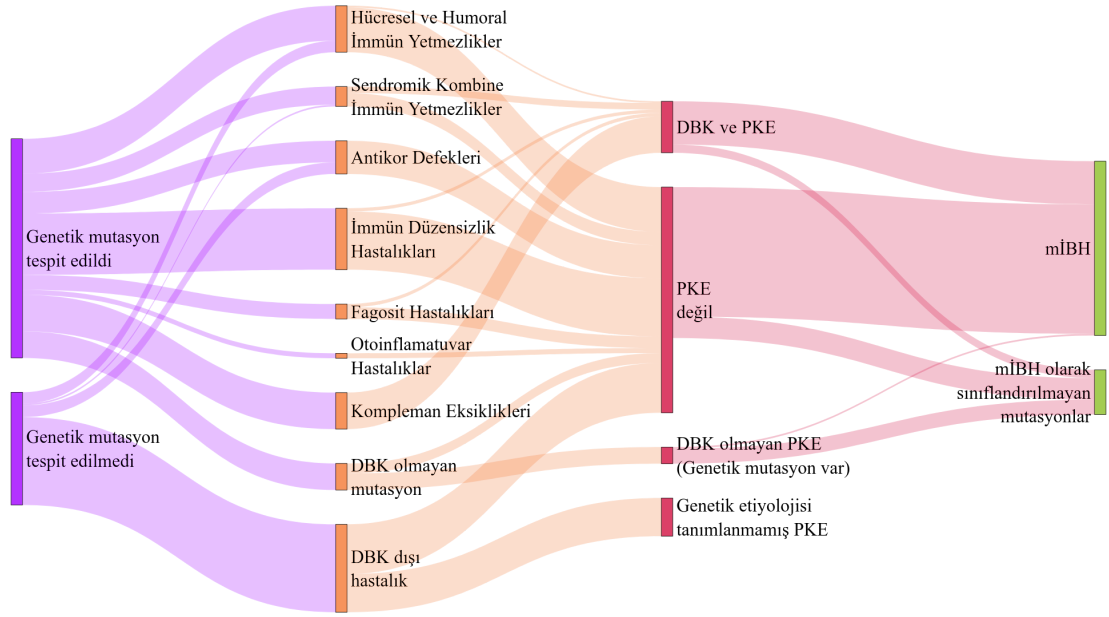
		Yüzde	Sayı
Cinsiyet	Kız	%36	72
	Erkek	%64	128
Ölüm	Yok	%82,6	147
	Var	%17,4	31
Akraba evliliği	Yok	%38,9	75
	Var	%61,1	118
Ailede bilinen bağırsak hastalığı	Yok	%78,7	140
	Var	%21,3	38
Ailede tanı konulmamış bağırsak semptomları (İshal, karın ağrısı vb)	Yok	%83,1	147
	Var	%16,9	30

Araştırma evrenine dahil edilen hastaların %66'sında (n = 132) genetik etyoloji tespit edilmiştir. Hastaların %65,5'inde (n = 131) güncel literatüre göre ⁴ DBK olarak sınıflanmış hastalık ve mutasyonlar, %8'inde (n = 16) DBK sınıflamasına dahil edilmemiş genetik mutasyonlar saptanmış, %26,5'inin (n = 53) ise genetik etyolojisi belirlenememiş ve DBK sınıflaması dışında hastalıklar olarak tespit edilmiştir. Hastaların %68'inde (n = 136) PKE şüphesi mevcuttur. Genetik etyolojisi tespit edilen hastaların %79,5'inde (n = 105), güncel literatüre göre ¹, mİBH olarak

sınıflandırılan bir gende mutasyon saptanmıştır. Araştırma evrenine dahil edilen hastaların genetik mutasyonlarının sıklıkları Tablo 6’da sunulmuş, farklı sınıflama sistemlerine göre dağılımları ise Tablo 7’de ve Şekil 8’de özetlenmiştir.

Tablo 6: Araştırma evrenindeki hastalarda görülen genetik mutasyonlar ve sıklıkları

Gen	Frekans	Gen	Frekans	Gen	Frekans
<i>CD55</i>	%11 (22)	<i>MALT1</i>	%1 (2)	<i>TYMP</i>	%0,5 (1)
<i>LRBA</i>	%7 (14)	<i>TTC7A</i>	%0,5 (1)	<i>TNFRSF13B</i>	%0,5 (1)
<i>CTLA4</i>	%4,5 (9)	<i>DPAGT1</i>	%0,5 (1)	<i>IGHM</i>	%0,5 (1)
<i>RAG1</i>	%3,5 (7)	<i>TTC37</i>	%0,5 (1)	<i>SPI10</i>	%0,5 (1)
<i>BTK</i>	%3 (6)	<i>SLCO2A1</i>	%0,5 (1)	<i>ALG6</i>	%0,5 (1)
<i>FOXP3</i>	%2,5 (5)	<i>XIAP</i>	%0,5 (1)	<i>ST6GALNAC1</i>	%0,5 (1)
<i>CARMIL2</i>	%1,5 (3)	<i>PIEZO1</i>	%0,5 (1)	<i>IL37</i>	%0,5 (1)
<i>DGAT1</i>	%1,5 (3)	<i>NCF4</i>	%0,5 (1)	<i>POMP</i>	%0,5 (1)
<i>CD40LG</i>	%1,5 (3)	<i>CYBA</i>	%0,5 (1)	<i>PRKCD</i>	%0,5 (1)
<i>MEFV</i>	%1,5 (3)	<i>MPI</i>	%0,5 (1)	<i>NCF2</i>	%0,5 (1)
<i>PIK3CD</i>	%1,5 (3)	<i>FAT4</i>	%0,5 (1)	<i>IL10RA</i>	%0,5 (1)
<i>CYBB</i>	%1,5 (3)	<i>PEPD</i>	%0,5 (1)	<i>DCLRE1C</i>	%0,5 (1)
<i>G6PC3</i>	%1,5 (3)	<i>PIK3R1</i>	%0,5 (1)	<i>IL2RG</i>	%0,5 (1)
<i>ADA</i>	%1,5 (3)	<i>ICOS</i>	%0,5 (1)	<i>IL2RA</i>	%0,5 (1)
<i>ADAMTS3</i>	%1 (2)	<i>RAG2</i>	%0,5 (1)	<i>DNMT3B</i>	%0,5 (1)
<i>CCBE1</i>	%1 (2)	<i>ZBTB24</i>	%0,5 (1)	Mutasyon tespit edilemedi	%34 (68)
<i>ARPC1B</i>	%1 (2)	<i>CD19</i>	%0,5 (1)		
<i>PIK3C2A</i>	%1 (2)	<i>STXBP2</i>	%0,5 (1)		
<i>ZAP70</i>	%1 (2)	<i>PGM3</i>	%0,5 (1)		



Şekil 8: Araştırma evrenindeki hastaların genetik etyolojilerinin farklı sınıflamalara göre dağılımı

Eflatun ile gösterilen çizgi dağılımı genetik mutasyon bulunan ve bulunmayan olguları göstermekteyken turuncu çizgiler DBK sınıflamasına karşılık gelen genetik bozuklukları ve hastalıkları sunmaktadır. Pembe çizgiler PKE ve DBK kategorilerini gösterirken yeşil çizgiler mİBH sınıflamasına karşılık gelen hasta dağılımını sunmaktadır.

Hastaların semptom başlangıç yaşı, boy ve kilo z skorlarının ortanca değerleri sırasıyla 9 ay (IQR 2 - 36), -1,4 (IQR -6,72 - 3,56) ve -1,84 (IQR -9,79 - 2,09) olarak bulunmuştur. Araştırma evreninde yer alan tüm hastaların semptom başlangıç yaşı ile boy ve kilo z skoru dağılımları Şekil 8’de sunulmuştur. Hastalıklar geniş bir yaş aralığında başlayabilmesine rağmen, büyük çoğunluğu hayatın ilk birkaç yılı içinde belirtiler göstermektedir. Bu hastalarda boy ve kilo değerlerinin toplum normlarına göre daha düşük olduğu gözlemlenmiştir. Yapılan analizler, boy uzunluğu üzerindeki etkinin kilo alımına kıyasla daha az olduğunu ortaya koymaktadır.

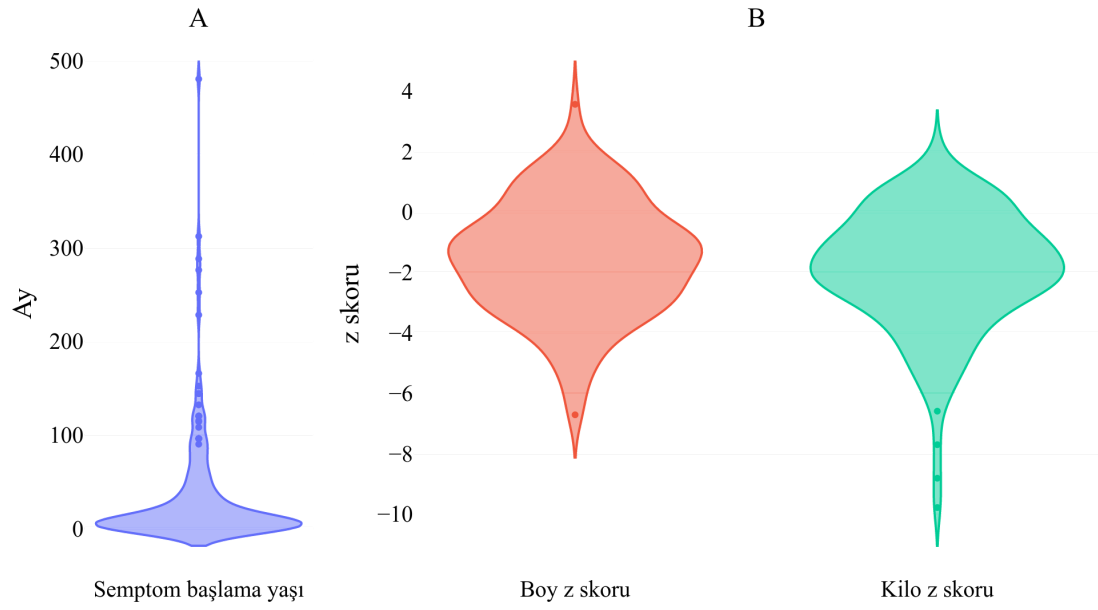
Semptom başlangıç yaşı, boy z skoru ve kilo z skorunun DBK sınıflamasına göre dağılımları Şekil 10’da ve Şekil 11’de gösterilmiştir. Analizler sonucunda, DBK sınıfları arasında boy z skoru farkının istatistiksel olarak anlamlı olduğu tespit edilmiştir ($p=0,036$). Ancak, yapılan ikili karşılaştırmalar neticesinde sınıflar

arasında anlamlı bir farkın saptanmadığı görülmüştür. Kilo z skorları ($p=0,068$) ve semptom başlama yaşı ($p = 0,056$) için ise gruplar arasında istatistiksel olarak anlamlı bir fark bulunmamıştır. Yani, boy ve kilo z skorları ile semptom başlama yaşı, hastalık gruplarını birbirinden ayırmada yetersiz kalmaktadır.

Semptom başlangıç yaşı, boy z skoru ve kilo z skorunun PKE gruplarına göre dağılımları Şekil 12’de ve Şekil 13’de görülmektedir. PKE grupları arasında yapılan analizler sonucunda, boy z skoru ($p=0,116$), kilo z skoru ($p = 0,364$) ve semptom başlama yaşının ($p = 0,480$) gruplar arasında istatistiksel olarak anlamlı bir fark taşımadığı tespit edilmiştir. Bu sonuçlar, boy ve kilo z skorlarının yanı sıra semptom başlama yaşının, hastalık gruplarını birbirinden ayırmada yetersiz olduğunu göstermektedir.

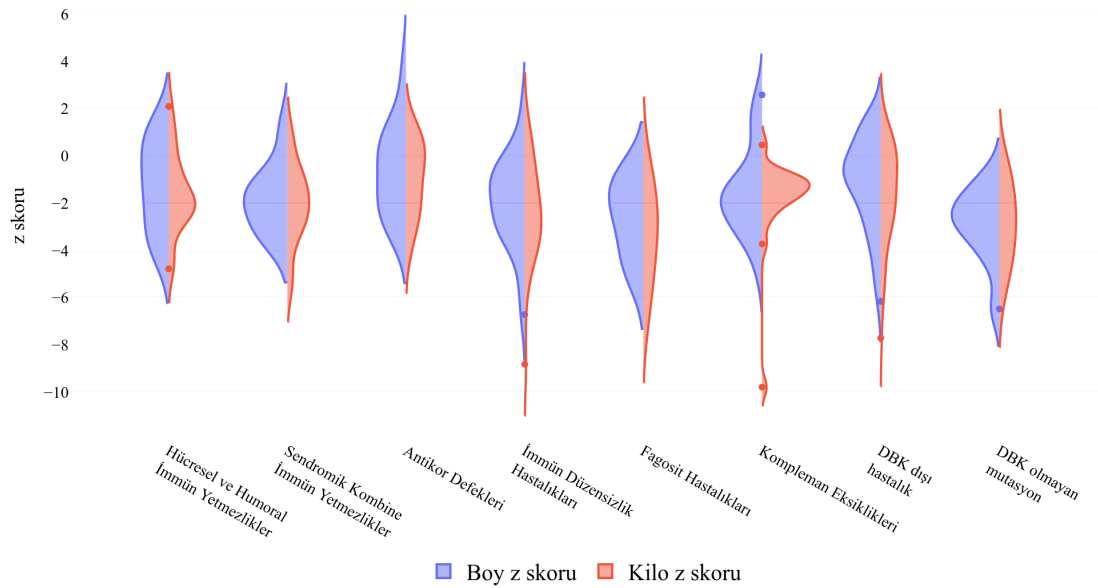
Tablo 7: Araştırma evrenindeki hastaların genetik mutasyonlarının farklı sınıflamalara göre dağılımı

		Yüzde	Sayı
DBK sınıflaması	Hüresel ve Humoral immün yetmezlikler	%14	28
	Sendromik kombine immün yetmezlikler	%6	12
	Antikor defekleri	%10	20
	İmmün düzensizlik hastalıkları	%18,5	37
	Fagosit hastalıkları	%4,5	9
	Otoinflamatuvar hastalıklar	%1,5	3
	Kompleman Eksiklikleri	%11	22
	DBK olmayan mutasyon	%8	16
	DBK dışı hastalık (Mutasyon tespit edilemedi)	%26,5	53
PKE gruplaması	DBK ve PKE	%15,5	31
	DBK olmayan PKE (Genetik mutasyon var)	%5	10
	Genetik etyolojisi tanımlanamamış PKE	%11,5	23
	PKE değil	%68	136
mİBH gruplaması	mİBH ilişkili	%79,5	105
	mİBH ilişkili değil	%20,5	27



Şekil 9: Semptom başlama yaşı, boy z skoru ve kilo z skoru dağılımları

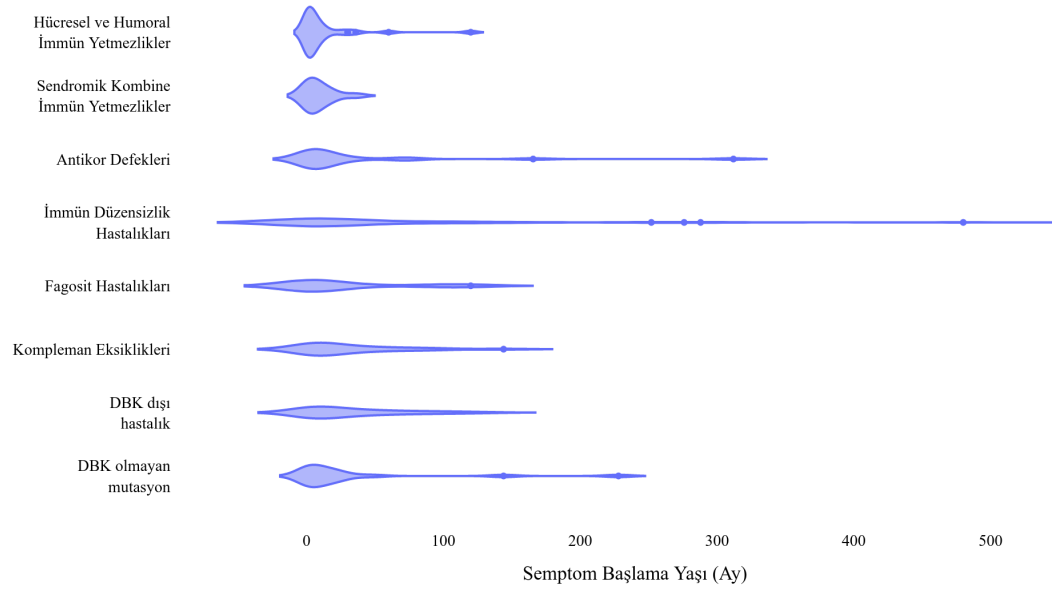
Boy z skorları normal dağılım göstermektedir ($p=0,739$) ancak semptom başlangıç yaşı ($p<0,001$) ve kilo z skorları normal dağılım göstermemektedir ($p<0,001$). Nokta ile gösterilen değerler aykırı değerlerdir.



Şekil 10: DBK sınıflamasına göre ayrılan gruplarda boy ve kilo z skoru dağılımı

DBK dışı hastalık grubunda boy ve kilo z skorlarının ($p=0,031$ ve $p=0,01$), Kompleman Eksiklikleri grubunda kilo z skorunu ($p<0,001$) normal dağılım

göstermemektedir. Otoinflamatuvar hastalık grubunda yeterli veri olmadığı için çıkarılmıştır. Nokta ile gösterilen değerler aykırı değerlerdir.



Şekil 11: DBK sınıflamasına göre ayrılan gruplarda semptom başlama yaşı dağılımı
Tüm gruplarda semptom başlama yaşı normal dağılım göstermemektedir.
Otoinflamatuvar hastalık grubunda yeterli veri olmadığı için çıkarılmıştır. Nokta ile gösterilen değerler aykırı değerlerdir.

Tablo 8: DBK sınıflamasına göre ayrılan gruplarda boy z skoru, kilo z skoru ve semptom başlangıç yaşı dağılımı

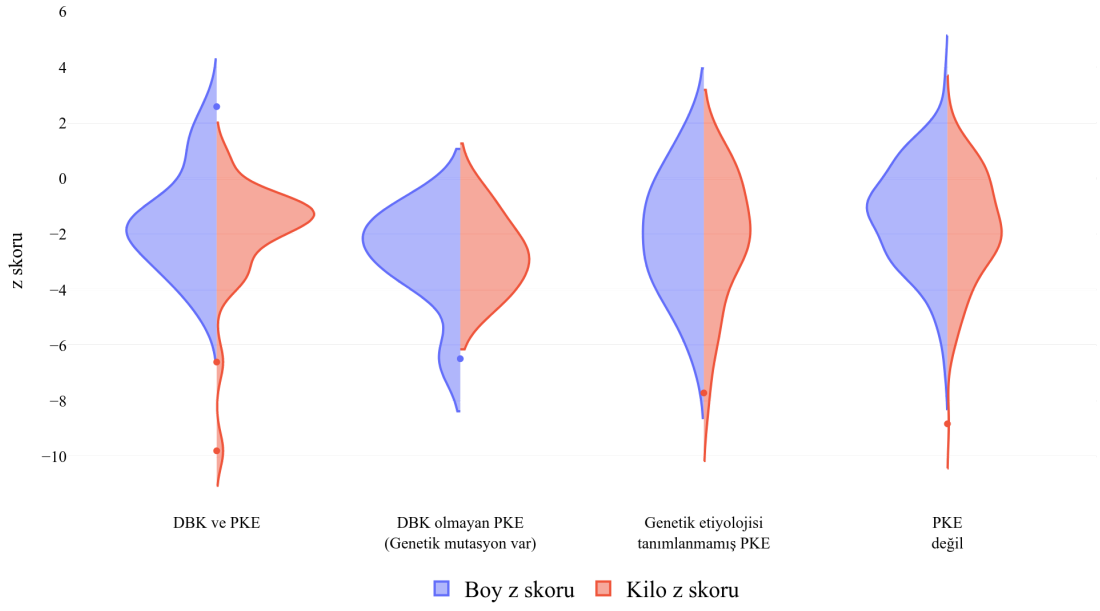
	Boy	Kilo	Semptom başlama yaşı
	Ortalama (SS)	Ortalama (SS)	Ortanca (IQR)
Hücrel ve Humoral immün yetmezlikler	-1,28 (1,87)	-1,59 (1,69)	4,5 (1 - 12)
Sendromik kombine immün yetmezlikler	-1,63 (1,38)	-1,65 (1,31)	6 (0 - 12,75)
Antikor defekleri	-0,66 (1,9)	-0,97 (1,57)	10,5 (2,5 - 27)
İmmün düzensizlik hastalıkları	-1,98 (1,88)	-2,45 (2,2)	8 (2 - 84)
Fagosit hastalıkları	-2,6 (1,62)	-3,34 (2,07)	3 (2 - 30)
Kompleman Eksiklikleri	-1,54 (1,87)	-1,88 (2,0)	12 (6,25 - 46,5)
DBK dışı hastalık (mutasyon tespit edilemedi)	-1,14 (1,92)	-1,66 (2,07)	21 (4,75 - 55,5)
DBK olmayan mutasyon	-2,91 (1,57)	-2,85 (1,73)	9 (2,75 - 24)

SS: Standart Sapma, IQR: Interquartile Range (Çeyrekler Arası Aralık)

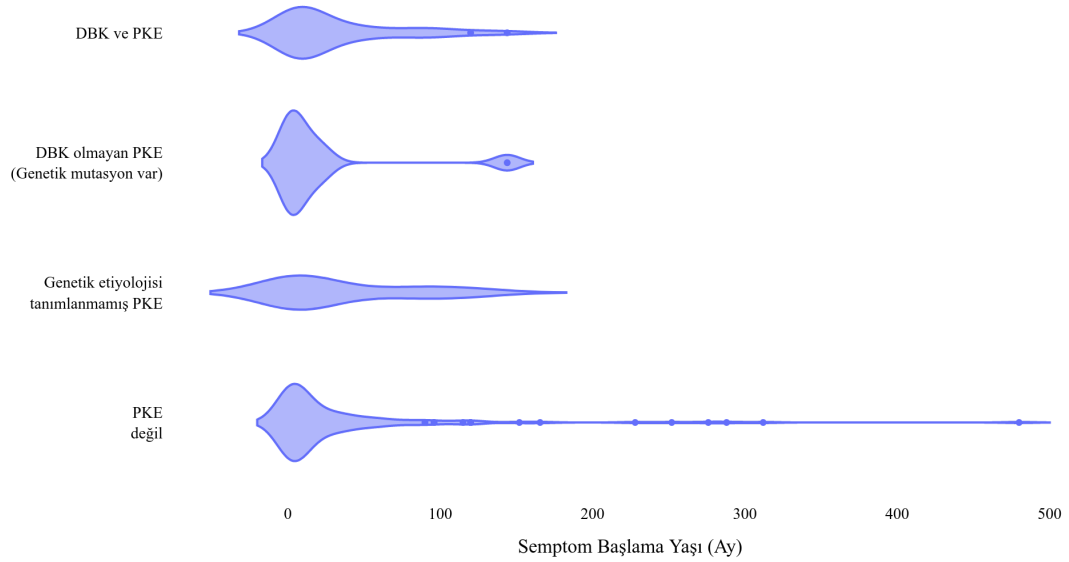
Tablo 9: PKE gruplamasına göre ayrılan gruplarda boy z skoru, kilo z skoru ve semptom başlangıç yaşı dağılımı

	Boy	Kilo	Semptom başlama yaşı
	Ortalama (SS)	Ortanca (IQR)	Ortanca (IQR)
DBK ve PKE	-1,57 (1,84)	-1,51 (-2,5 - -1,05)	12,0 (6,0 - 45,0)
DBK olmayan PKE (Genetik mutasyon var)	-2,78 (1,77)	-2,66 (-3,6 - -2,14)	3,3 (2,25 - 17,25)
Genetik etyolojisi tanımlanmamış PKE	-2,01 (2,11)	-2,12 (-2,84 - -0,11)	11,0 (2,75 - 73,5)
PKE değil	-1,32 (1,84)	-1,83 (-2,8 - -0,3)	7,0 (2,0 - 36,0)

SS: Standart Sapma, IQR: Interquartile Range (Çeyrekler Arası Aralık)

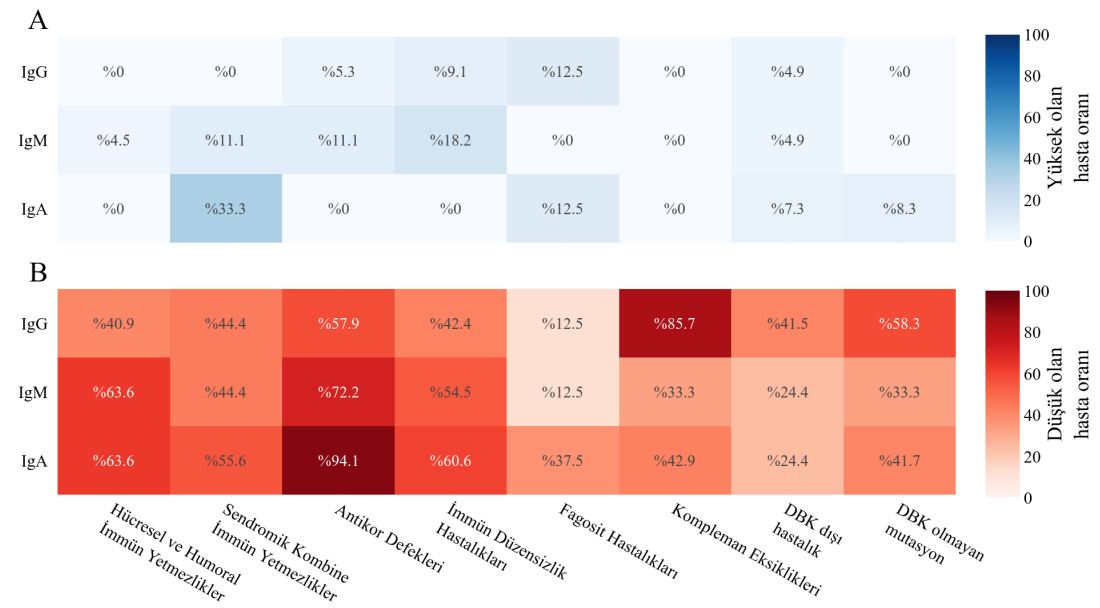


Şekil 12: PKE gruplamasına göre ayrılan gruplarda boy ve kilo z skoru dağılımı
Boy z skoru tüm gruplarda; kilo z skoru DBK olmayan PKE ($p=0,578$) ve genetik etiyolojisi tanımlanmamış PKE ($p=0,108$) gruplarında normal dağılım göstermektedir. Nokta ile gösterilen değerler aykırı değerlerdir.



Şekil 13: DBK sınıflamasına göre ayrılan gruplarda semptom başlama yaşı dağılımı
Semptom başlama yaşının, hiçbir grupta normal dağılım göstermemektedir. Nokta ile gösterilen değerler aykırı değerlerdir.

Hastaların yaşına göre değerlendirilen Ig seviyelerinin DBK sınıfları arasındaki farka baktığımızda IgG ($p=0,08$) ve IgA'nın ($p<0,001$) düzeylerinin anlamlı olarak farklılık gösterdiği gözlenmiştir (Şekil 14). İlgili analiz sonuçlarına göre, IgA seviyeleri özellikle Antikor Defektleri sınıfında, IgG seviyeleri ise özellikle Kompleman Hastalıkları sınıfında önemli derecede düşük bulunmuştur. Fagosit Hastalıklar sınıfında ise, diğer sınıfların aksine, Ig seviyeleri genellikle normal ya da yüksek seviyelerde bulunmaktadır.



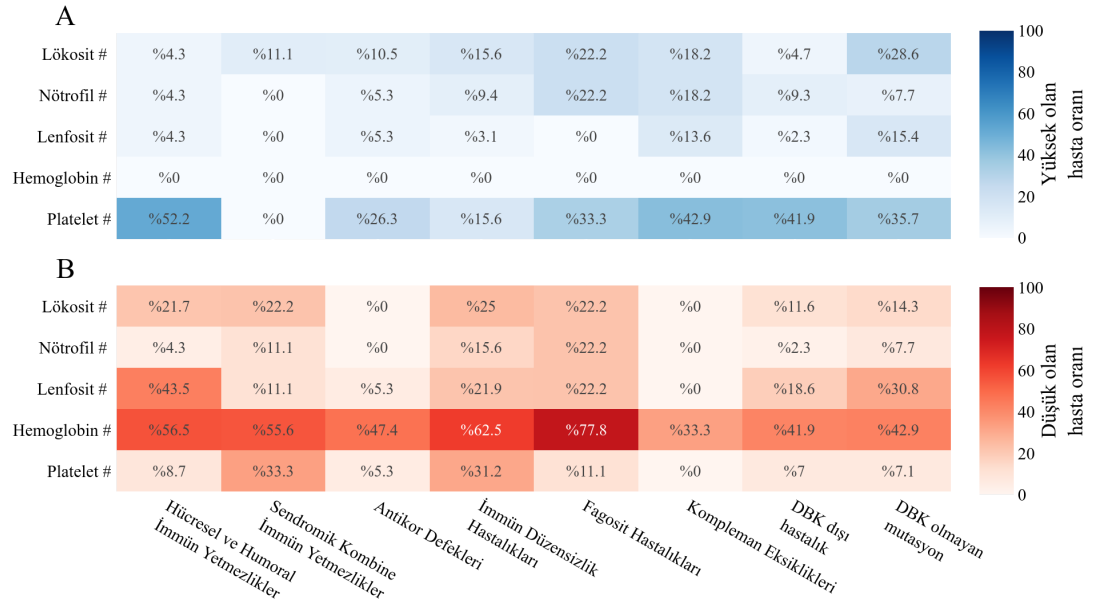
Şekil 14: Ig seviyelerinin yaşa göre durumunun DBK sınıflamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

Hastaların yaşına göre değerlendirilen hemogram değerlerinin DBK sınıfları arasındaki farka baktığımızda, lenfosit sayısı ($p=0,021$) ve trombosit sayısının ($p=0,002$) gruplar arasında anlamlı fark gösterdiği tespit edilmiştir (Şekil 15). Hemogloblin değeri ise, özellikle Fagosit Hastalıklar ve İmmün Düzensizlik Hastalıkları gruplarında olmak üzere, tüm gruplarda düşük bulunmaktadır. Trombosit sayısı ise Hücrel ve Humoral İmmün Yetmezlikler, Kompleman Hastalıkları ve

İmmün Eksiklik Hastalıkları dışında yüksek seviyelerde görülmektedir. Her ne kadar ikili karşılaştırmalarda anlamlı farklar bulunmasa da, lenfosit sayısının özellikle Hücresel ve Humoral İmmün Yetmezlik hastalığında daha sık düşük bulunduğu gözlemlenmiştir.



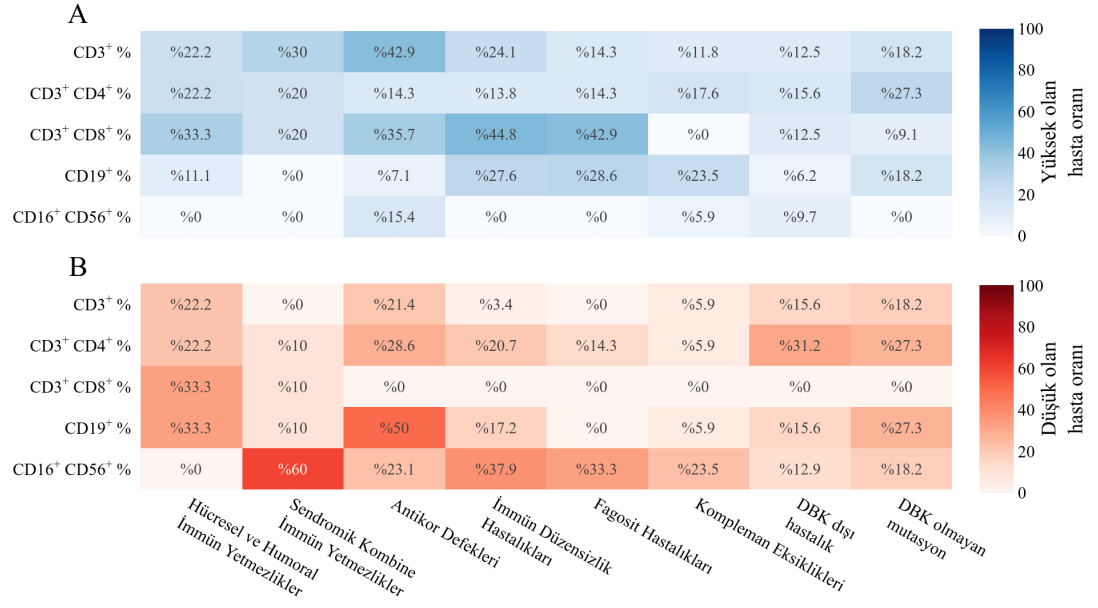
Şekil 15: Hemogramın yaşa göre durumunun DBK sınıflamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

Hastaların yaşa göre değerlendirilen lenfosit alt grupları seviyelerinin DBK sınıfları arasındaki farkına baktığımızda, CD3⁺ CD8⁺ (CD8 T Lenfosit) oranı (p=0,017), CD19⁺ (B Lenfosit) sayısı (p=0,031), CD16⁺ CD56⁺ (NK) oranı (p=0,028) ve CD16⁺ CD56⁺ sayısı (p=0,006) arasında anlamlı farklar tespit edilmiştir (Şekil 16 ve Şekil 17).

CD3⁺ (T Lenfosit) oranı genel olarak DBK sınıfları içinde normal düzeylerde bulunurken, sayısal düşüklüğü Kompleman Eksiklikleri dışında tüm sınıflarda belirgin bir özelliktir. DBK dışı hastalık ve DBK dışı mutasyon grupları, CD3⁺ sayısında belirgin bir eğilim göstermemektedir.



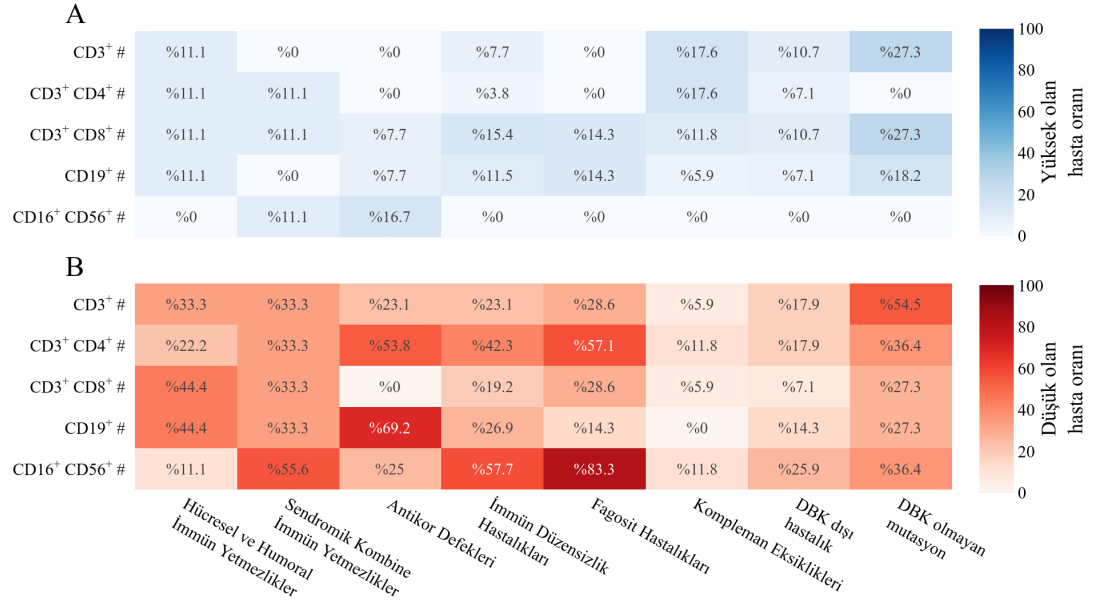
Şekil 16: Basit immünofenotiplemede hücre oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

CD3⁺ CD4⁺ (CD4⁺ T Lenfosit) oranı genellikle tüm sınıflarda normal düzeylerdedir, ancak Antikor Defektleri ve Fagosit Hastalıkları'nda sayısındaki düşüklük daha yaygın görülmektedir. CD3⁺ CD8⁺ oranı, İmmün Düzensizlik Hastalıkları ve Fagosit Hastalıklarında normal-yüksek seyredebilirken, düşüklüğü en sık Hücrel ve Humoral İmmün Yetmezlikler'de gözlenmektedir. Antikor Defektleri olan hastalarda ise CD3⁺ CD8⁺ oranı ve sayısı genellikle normal sınırlar içindedir.

Antikor Defektlerinde CD19⁺ (B Lenfosit) oranı ve sayısı genellikle düşük-normal seyretmektedir. CD16⁺ CD56⁺ (NK) oranı en sık Sendromik Kombine İmmün Yetmezlikler'de düşük olurken, sayısı ise en sık Fagositer Hastalıklar'da görülmektedir.



Şekil 17: Basit immünofenotiplemede hücre sayılarının yaşa göre durumunun DBK sınıflamasına göre dağılımı

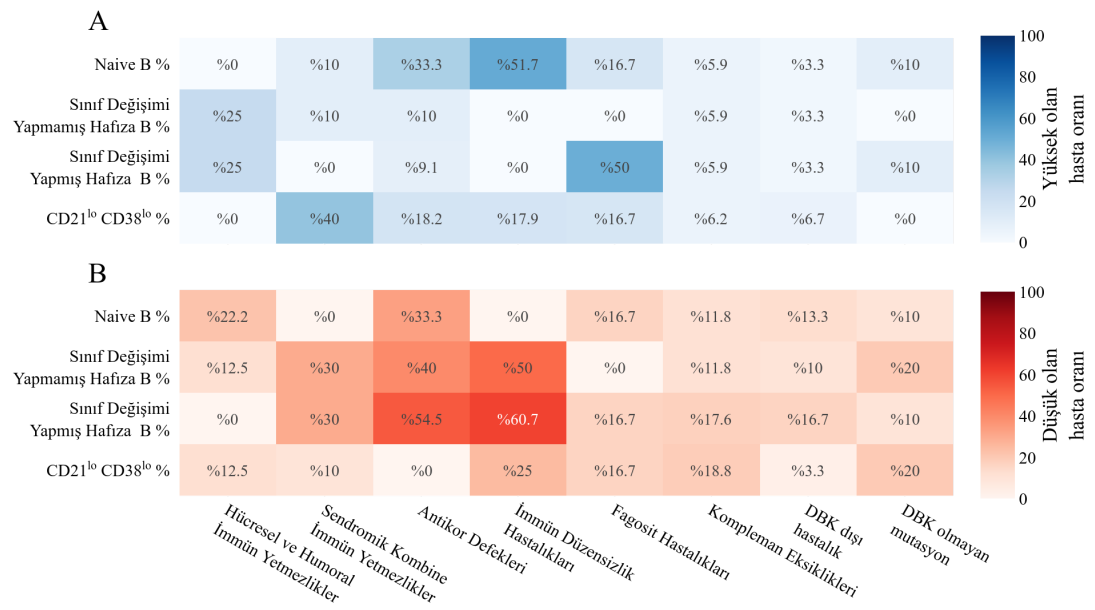
A: Değeri yaşına göre düşük olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B Lenfosit alt gruplarının DBK sınıfları arasındaki farkına baktığımızda, Naive B Lenfosit oranı ($p<0,001$) ve sayısının ($p=0,004$); sınıf değişimi yapmamış Hafıza B Lenfosit oranı ($p=0,012$) ve sayısının ($p=0,004$); sınıf değişimi yapmış Hafıza B Lenfosit oranı ($p<0,001$) ve sayısının ($p<0,001$) sınıflar arasında istatistiksel olarak anlamlı farklar gösterdiği tespit edilmiştir (Şekil 18, Şekil 19).

Antikor Defektleri'nde ve İmmün Düzensizlik Hastalıkları'nda, sınıf değişimi yapmış ve yapmamış olan B Lenfosit oranlarının genellikle düşük olduğu tespit edilmiştir. Hücre gruplarının sayısal değerleri, özellikle Antikor Defektleri olmak üzere birçok sınıfta düşük bulunmuştur. Ancak, bu durum Kompleman Eksiklikleri için geçerli değildir. Diğer hücre gruplarına kıyasla B Lenfosit alt gruplarının sayısal değerleri, DBK dışı mutasyonlarda da çoğunlukla normal ya da düşüktür.

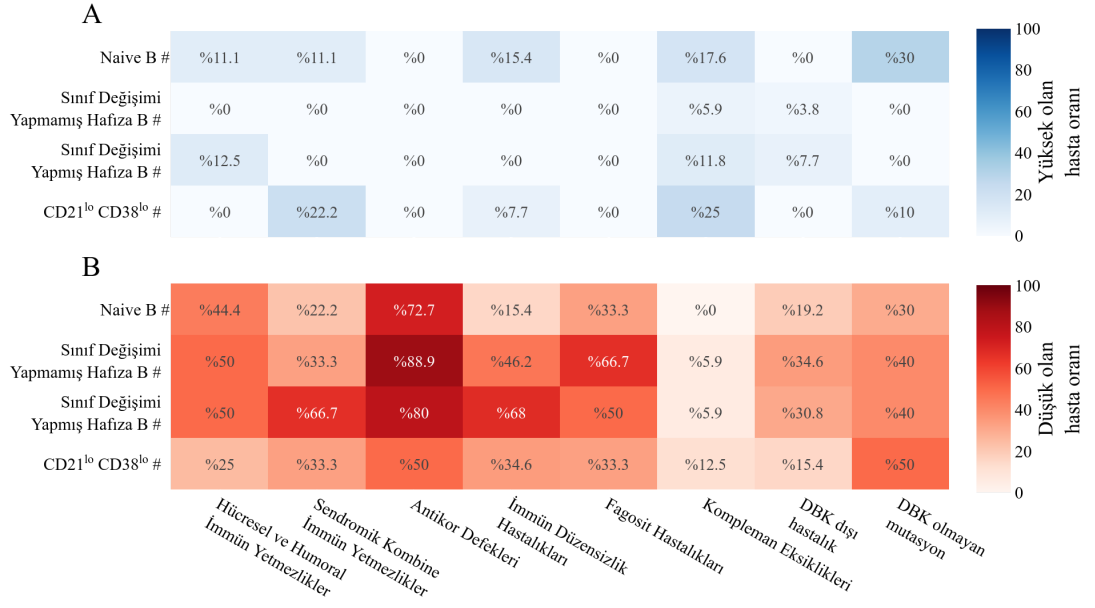
T Lenfosit alt gruplarının DBK sınıfları arasındaki farklara baktığımızda, CD8 Hafıza oranı ($p=0,006$) dışında hiçbir hücre grubunda istatistiksel anlamlı fark tespit edilmemiştir (Şekil 20, Şekil 21, Şekil 22) RTE oranı genellikle Hücresel ve Humoral İmmün Yetmezlikler ile DBK dışı hastalıklarda düşük ya da normal seviyelerde görülmektedir. Hafıza CD4 oranı ise Hücresel ve Humoral İmmün Yetmezlikler’de genelde yüksek seyreder. Sayısal değerlere göre bakıldığında Naive CD8 sayısının Fagosit Hastalıkları ve DBK dışı mutasyonlarda düşük olduğu görülmektedir.



Şekil 18: B Lenfosit alt gruplarının oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

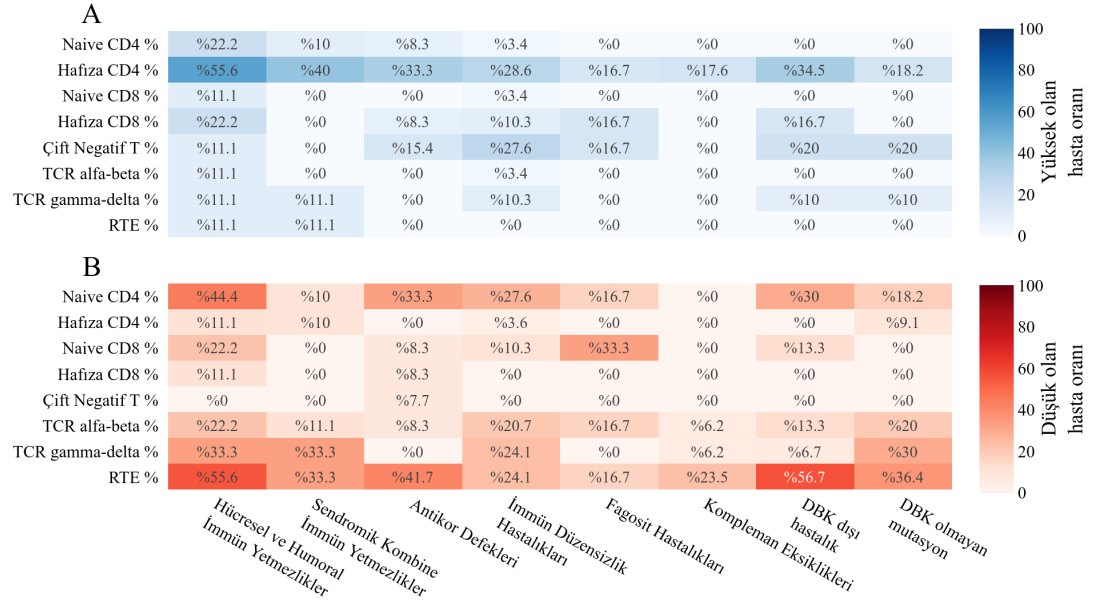


Şekil 19: B Lenfosit alt gruplarının sayılarının yaşa göre durumunun DBK sınıflamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayısının, aynı DBK sınıfındaki tüm hasta sayısına oranı

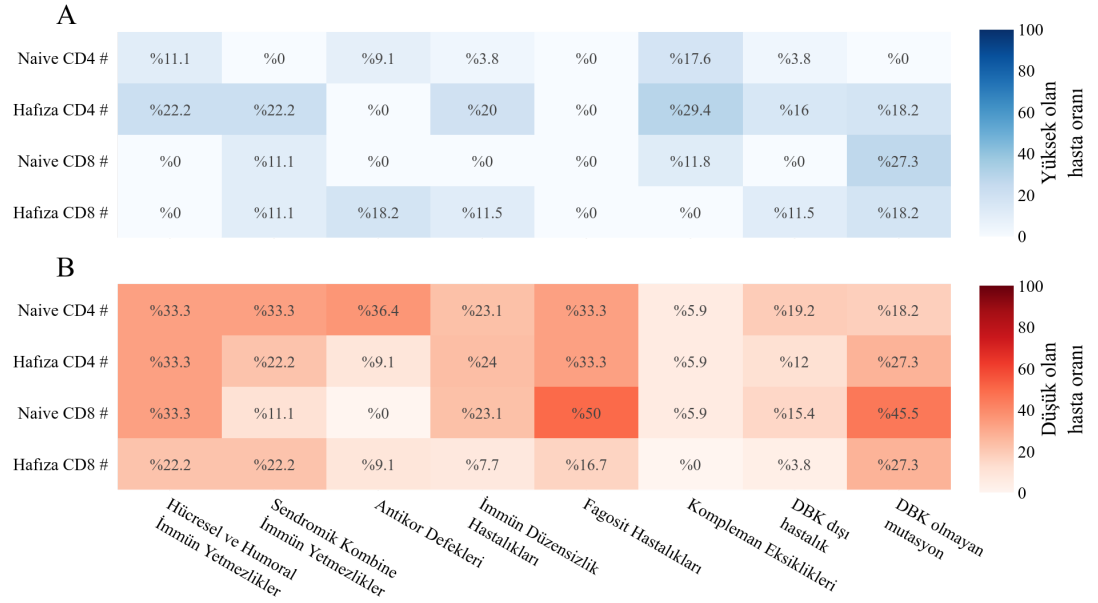
CD4 ve CD8 T Lenfosit alt gruplarında, CD4 Efektör Hafıza oranının genellikle tüm sınıflarda yüksek izlendiği gözlenmektedir. Ancak CD4 Naive ve CD4 Esas Hafıza oranları, tüm gruplarda düşük seyretmektedir. Ayrıca, CD8 Efektör Hafıza oranlarının özellikle Antikor Defektleri, Fagosit Hastalıkları ve DBK dışı hastalıklarda sıklıkla yüksek ölçüldüğü belirlenmiştir.



Şekil 20: T Lenfosit alt gruplarının oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

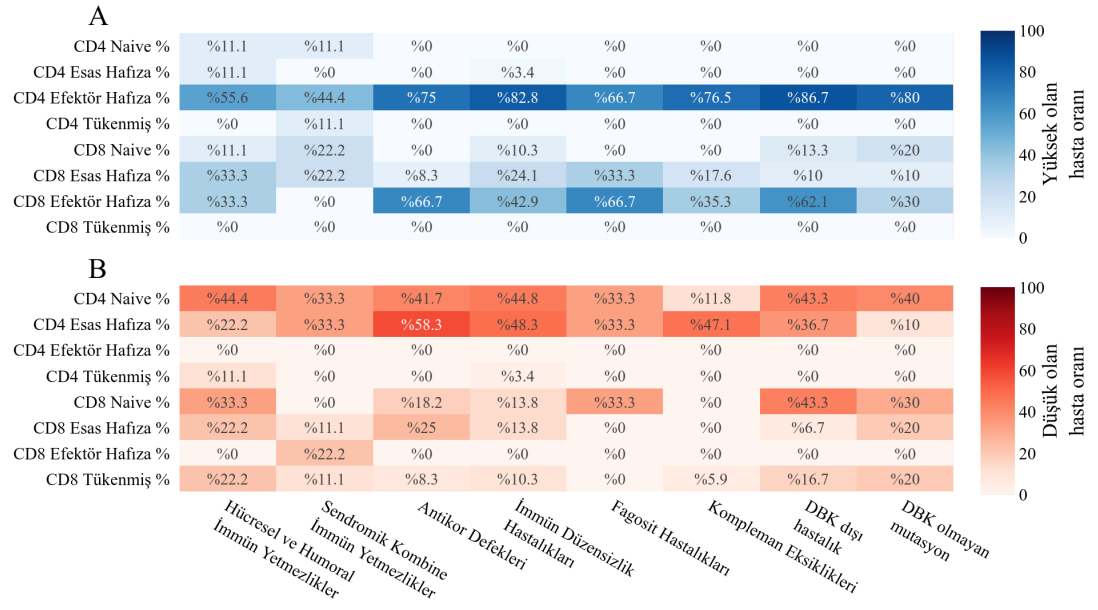
B: Değeri yaşına göre yüksek olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı



Şekil 21: T hücre alt gruplarının sayılarının yaşa göre durumunun DBK sınıflamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı



Şekil 22: CD4 ve CD8 T lenfosit alt gruplarının oranlarının yaşa göre durumunun DBK sınıflamasına göre dağılımı

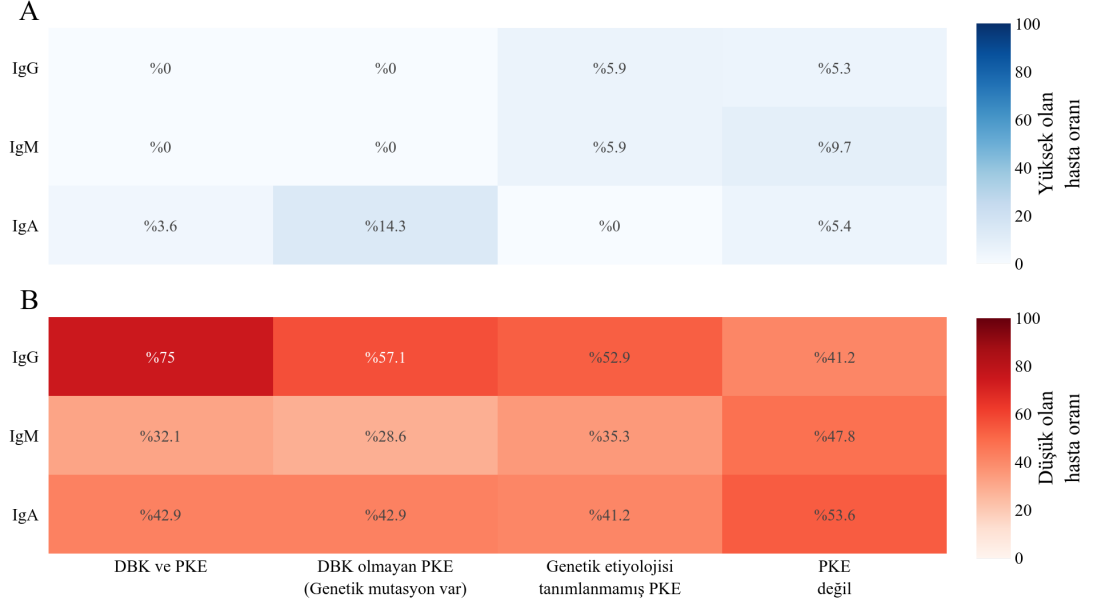
A: Değeri yaşına göre düşük olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

Hastaların yaşına göre değerlendirilen Ig seviyelerinin PKE grupları arasındaki farkına baktığımızda, sadece IgG seviyesinin (p=0,012) istatistiksel olarak anlamlı fark gösterdiği gözlenmiştir. Özellikle PKE ile seyreden DBK hastalarında diğer gruplara kıyasla IgG seviyelerinin daha sık olarak düşük olduğu belirlenmiştir. Ayrıca, grupların tamamında hastaların yaklaşık yarısında Ig düşüklüğüne sahip olduğu gözlenmektedir (Şekil 23).

Hastaların yaşına göre değerlendirilen hemogram değerlerinin PKE grupları arasındaki farkına baktığımızda, istatistiksel olarak anlamlı bir fark gözlenmediği

belirlenmiştir. Ancak, Hemogloblin seviyesinin en sık PKE dışı hastalıklarda düşük olduğu görülmektedir (Şekil 24).



Şekil 23: Ig seviyelerinin yaşa göre durumunun PKE gruplamasına göre dağılımı

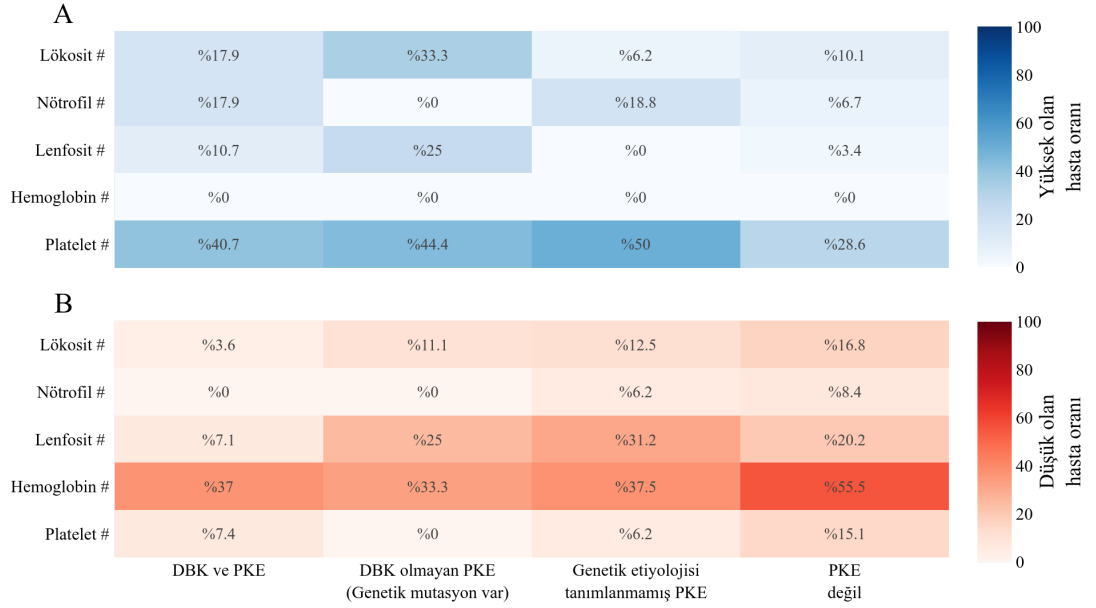
A: Değeri yaşına göre düşük olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı

Lenfosit alt gruplarına bakıldığında gruplar arasında anlamlı bir fark görülmedi. Ancak, $CD3^+ CD4^+$ oranı düşüklüğünün en sık, genetik etiyolojisi tanımlanmamış PKE grubunda bulunmaktadır (Şekil 25).

B Lenfosit alt gruplarından, sınıf değişimi yapmamış Hafıza B oranının ($p=0,039$) PKE grupları arasında farklılık gösterdiği tespit edilmiştir. Bu hücre grubu, PKE dışı hastalıklarda diğer gruplardan daha sık düşük gözlenmiştir (Şekil 26).

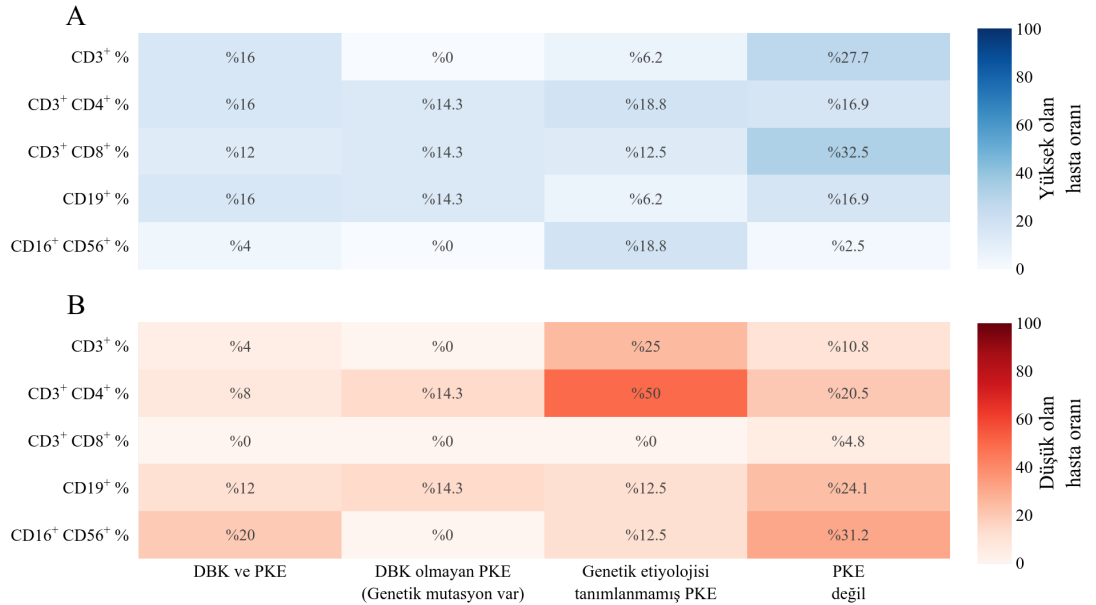
T Lenfosit alt gruplarından RTE ($p=0,037$), PKE grupları arasında istatistiksel olarak anlamlı fark göstermektedir. Özellikle genetik etiyolojisi tanımlanmamış PKE hastalarında RTE oranının sıklıkla düşük olduğu gözlenmiştir (Şekil 27).



Şekil 24: Hemogramın yaşa göre durumunun PKE gruplaması göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

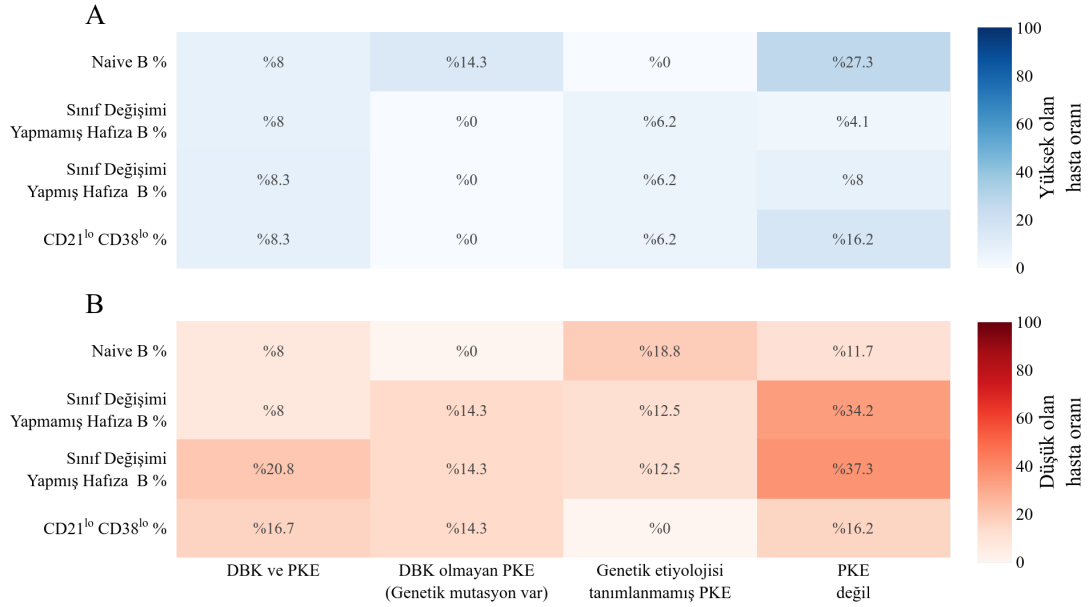


Şekil 25: Basit immünofenotiplemede hücre oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı DBK sınıfındaki tüm hasta sayısına oranı

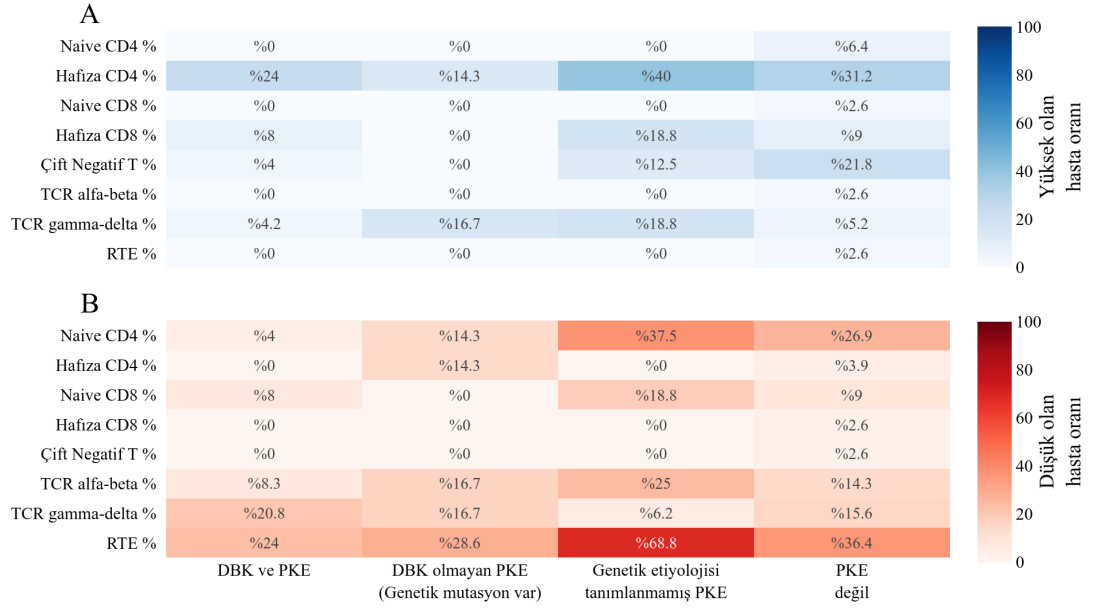
CD4 T Lenfosit alt gruplarının, PKE grupları arasında istatistiksel olarak anlamlı bir fark tespit edilememiştir. Ancak, CD4 Naive T Lenfositler, genetik etyolojisi belirlenememiş PKE grubunda sıklıkla düşük gözlenmektedir (Şekil 28). CD8 T Lenfosti alt gruplarında ise CD8 Naive oranının (p=0,016) gruplar arası anlamlı fark gösterdiği ve en sık genetik etyolojisi tanımlanmamış PKE hastalarında düşük olduğu gözlenmiştir. CD4 Efektör Hafıza T Lenfositleri'nin de grupların tamamında normalden yüksek seyrettiği görülmüştür (Şekil 28).



Şekil 26: B Lenfosit alt gruplarının oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı

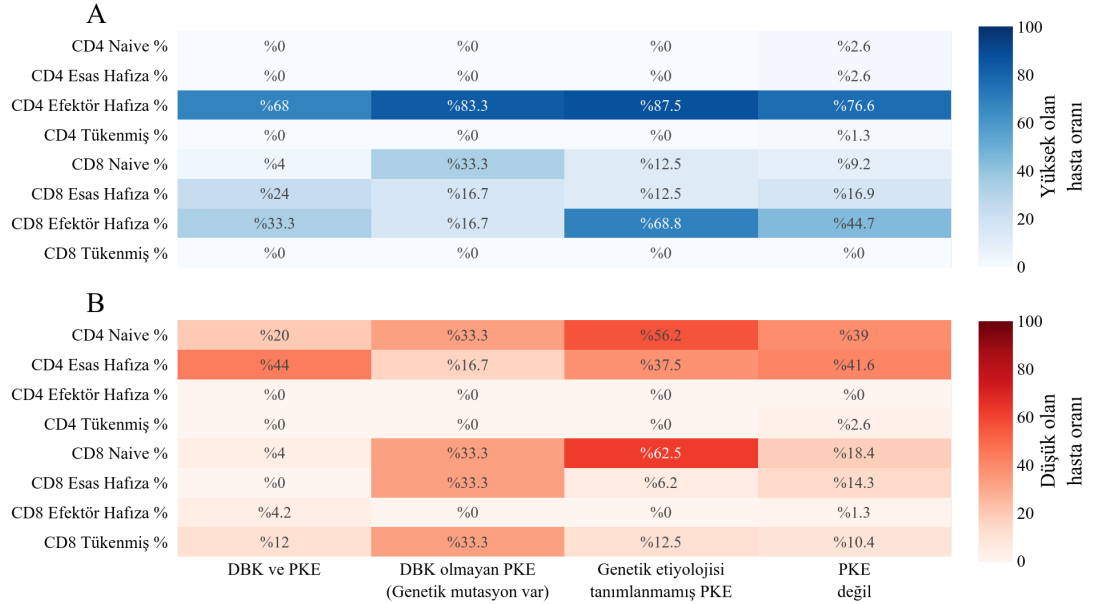
B: Değeri yaşına göre yüksek olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı



Şekil 27: T hücre alt gruplarının oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı

A: Değeri yaşına göre düşük olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı

B: Değeri yaşına göre yüksek olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı



Şekil 28: CD4 ve CD8 alt gruplarının oranlarının yaşa göre durumunun PKE gruplamasına göre dağılımı

A: Deęeri yaşına göre düşük olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı

B: Deęeri yaşına göre yüksek olan hasta sayının, aynı PKE grubundaki tüm hasta sayısına oranı

Araştırma evrenine dahil olan hastalarda toplamda 481 farklı HPO terimi tespit edilmiştir (Tablo 10). Dahil etme kriterlerinden olan Mukoid İshal (HP:0033343), Kronik İshal (HP:0002028), İshal (HP:0002014), Kanlı İshal (HP:0025085), İntestinal Lenfanjektazi (HP:0002593), tanımlı olan hastalardan silindi.

Hastaların başlangıç semptom ve bulgularının sıklığının ortancası 1 (IQR 1-3)'dir. Bu terimlerin sıklık dağılımı ve en sık gözlenen 10 HPO terimi Şekil 29'de gösterilmiştir. Hastaların, hastalığın seyri boyunca (başlangıçtan tanı konulana ve de varsa tedavi başlayana kadar) gözlenen semptom ve bulgularında ise ortanca 2 (IQR 1-7)'dir. Sıklık dağılımı ve en sık gözlenen 20 HPO terimi Şekil 30'da gösterilmiştir.

Hasta başına başlangıç semptom ve bulgu sayısı ortancası 1 (IQR 1 - 2) olarak hesaplanmıştır. Hasta başına, hastalık seyri boyunca tespit edilen semptom ve bulgu sayısı ortancası ise 19 (IQR 14 - 24) olarak hesaplanmıştır. Hasta başına tespit edilen HPO terimlerinin sıklıkları Şekil 30'da gösterilmektedir.

Tablo 10: Araştırma evrenindeki tüm HPO terimleri ve görülme sıklıkları

HPO	Tanım	Yüzde (Sayı)	HPO	Tanım	Yüzde (Sayı)
HP:0001903	Anaemia	%28.27 (136)	HP:0002028	Chronic diarrhea	%26.61 (128)
HP:0100512	Low levels of vitamin D	%20.58 (99)	HP:0003073	Hypoalbuminemia	%20.37 (98)
HP:0003075	Hypoproteinemia	%19.54 (94)	HP:0004315	Decreased circulating IgG level	%17.88 (86)
HP:0002720	Decreased circulating IgA level	%17.67 (85)	HP:0004325	Decreased body weight	%17.67 (85)
HP:0002014	Diarrhea	%16.63 (80)	HP:0002240	Hepatomegaly	%14.97 (72)
HP:0002850	Decreased circulating total IgM	%14.76 (71)	HP:0004322	Short stature	%14.35 (69)
HP:0002910	Elevated hepatic transaminase	%13.93 (67)	HP:0002027	Abdominal pain	%13.1 (63)
HP:0100502	Vitamin B12 deficiency	%12.89 (62)	HP:0001888	Lymphopenia	%12.89 (62)
HP:0002013	Vomiting	%12.68 (61)	HP:0001894	Thrombocytosis	%12.47 (60)
HP:0001744	Splenomegaly	%11.85 (57)	HP:0001945	Fever	%11.23 (54)
HP:0003270	Abdominal distention	%11.23 (54)	HP:0000964	Eczema	%10.4 (50)
HP:0001891	Iron deficiency anemia	%9.98 (48)	HP:0002593	Intestinal lymphangiectasia	%9.56 (46)
HP:0004313	Decreased circulating antibody level	%9.56 (46)	HP:0025085	Bloody diarrhea	%9.15 (44)
HP:0001510	Growth delay	%9.15 (44)	HP:0001508	Failure to thrive	%8.52 (41)
HP:0002090	Pneumonia	%8.32 (40)	HP:0006532	Recurrent pneumonia	%8.32 (40)
HP:0033164	Focal active colitis	%8.32 (40)	HP:0002039	Anorexia	%8.32 (40)
HP:0002716	Lymphadenopathy	%8.32 (40)	HP:0000282	Facial edema	%8.11 (39)
HP:0001873	Thrombocytopenia	%7.48 (36)	HP:0012735	Cough	%7.28 (35)
HP:0005263	Gastritis	%7.07 (34)	HP:0009098	Chronic oral candidiasis	%6.86 (33)
HP:0011227	Elevated circulating C-reactive protein concentration	%6.86 (33)	HP:0002018	Nausea	%6.65 (32)
HP:0001217	Clubbing	%6.44 (31)	HP:0001974	Leukocytosis	%6.24 (30)
HP:0001882	Leukopenia	%6.03 (29)	HP:0000969	Edema	%6.03 (29)
HP:0003077	Hyperlipidemia	%6.03 (29)	HP:0002719	Recurrent infections	%5.82 (28)
HP:0002788	Recurrent upper respiratory tract infections	%5.82 (28)	HP:0011107	Recurrent aphthous stomatitis	%5.82 (28)
HP:0000403	Recurrent otitis media	%5.82 (28)	HP:0012758	Neurodevelopmental delay	%5.61 (27)
HP:0002110	Bronchiectasis	%5.61 (27)	HP:0033343	Mucoid diarrhea	%5.2 (25)
HP:0002583	Colitis	%5.2 (25)	HP:0100827	Lymphocytosis	%4.99 (24)
HP:0000821	Hypothyroidism	%4.78 (23)	HP:0001875	Neutropenia	%4.78 (23)
HP:0000388	Otitis media	%4.57 (22)	HP:0002901	Hypocalcemia	%4.37 (21)
HP:0001047	Atopic dermatitis	%4.37 (21)	HP:0008151	Prolonged prothrombin time	%4.37 (21)
HP:0011897	Neutrophilia	%4.37 (21)	HP:0000246	Sinusitis	%4.16 (20)
HP:0003565	Elevated erythrocyte sedimentation rate	%4.16 (20)	HP:0001369	Arthritis	%3.95 (19)
HP:0002094	Dyspnea	%3.74 (18)	HP:0010741	Pedal edema	%3.74 (18)
HP:0002573	Hematochezia	%3.74 (18)	HP:0004936	Venous thrombosis	%3.53 (17)
HP:0025289	Cervical lymphadenopathy	%3.12 (15)	HP:0002099	Asthma	%2.91 (14)
HP:0012378	Fatigue	%2.91 (14)	HP:0003496	Increased circulating IgM level	%2.91 (14)
HP:0002829	Arthralgia	%2.91 (14)	HP:0003212	Increased circulating IgE level	%2.7 (13)

Tablo 10 (Devamı):

HPO	Tanım	Yüzde (Sayı)	HPO	Tanım	Yüzde (Sayı)
HP:0002205	Recurrent respiratory infections	%2.7 (13)	HP:0005225	Intestinal edema	%2.7 (13)
HP:0100750	Atelectasis	%2.7 (13)	HP:0002902	Hyponatremia	%2.49 (12)
HP:0031292	Cutaneous abscess	%2.49 (12)	HP:0001973	Autoimmune thrombocytopenia	%2.49 (12)
HP:0001999	Abnormal facial shape	%2.29 (11)	HP:0001954	Recurrent fever	%2.29 (11)
HP:0008940	Generalized lymphadenopathy	%2.29 (11)	HP:0100721	Mediastinal lymphadenopathy	%2.29 (11)
HP:0000365	Hearing impairment	%2.29 (11)	HP:0002783	Recurrent lower respiratory tract infections	%2.29 (11)
HP:0034274	Gastrointestinal ulcer	%2.29 (11)	HP:0100280	Crohn's disease	%2.29 (11)
HP:0001250	Seizure	%2.29 (11)	HP:0025615	Abscess	%2.29 (11)
HP:0500093	Food allergy	%2.29 (11)	HP:0000939	Osteoporosis	%2.29 (11)
HP:0002960	Autoimmunity	%2.08 (10)	HP:0100279	Ulcerative colitis	%2.08 (10)
HP:0025179	Ground-glass opacification	%2.08 (10)	HP:0001541	Ascites	%2.08 (10)
HP:0002917	Hypomagnesemia	%2.08 (10)	HP:0001878	Hemolytic anemia	%2.08 (10)
HP:0000988	Skin rash	%2.08 (10)	HP:0000083	Renal insufficiency	%2.08 (10)
HP:0001397	Hepatic steatosis	%2.08 (10)	HP:0001876	Pancytopenia	%2.08 (10)
HP:0033608	Pulmonary nodule	%2.08 (10)	HP:0000271	Abnormality of the face	%2.08 (10)
HP:0002728	Chronic mucocutaneous candidiasis	%2.08 (10)	HP:0100539	Periorbital edema	%2.08 (10)
HP:0030829	Abnormal breath sound	%2.08 (10)	HP:0011950	Bronchiolitis	%1.87 (9)
HP:0012387	Bronchitis	%1.87 (9)	HP:0002202	Pleural effusion	%1.66 (8)
HP:0003261	Increased circulating IgA level	%1.66 (8)	HP:0002020	Gastroesophageal reflux	%1.66 (8)
HP:0001880	Eosinophilia	%1.66 (8)	HP:0002900	Hypokalemia	%1.66 (8)
HP:0010783	Erythema	%1.66 (8)	HP:0100633	Esophagitis	%1.66 (8)
HP:0002019	Constipation	%1.66 (8)	HP:0002840	Lymphadenitis	%1.66 (8)
HP:0002315	Headache	%1.46 (7)	HP:0005523	Lymphoproliferative disorder	%1.46 (7)
HP:0003237	Increased circulating IgG level	%1.46 (7)	HP:0002148	Hypophosphatemia	%1.46 (7)
HP:0010447	Anal fistula	%1.46 (7)	HP:0001941	Acidosis	%1.46 (7)
HP:0033106	Elevated circulating D-dimer concentration	%1.46 (7)	HP:0001252	Hypotonia	%1.25 (6)
HP:0001747	Accessory spleen	%1.25 (6)	HP:0000787	Nephrolithiasis	%1.25 (6)
HP:0004396	Poor appetite	%1.25 (6)	HP:0033431	Cytomegalovirus colitis	%1.25 (6)
HP:0004432	Agammaglobulinemia	%1.25 (6)	HP:0002841	Recurrent fungal infections	%1.25 (6)
HP:0000093	Proteinuria	%1.25 (6)	HP:0001370	Rheumatoid arthritis	%1.25 (6)
HP:0033117	Duodenitis	%1.25 (6)	HP:0005353	Recurrent herpes	%1.25 (6)
HP:0001045	Vitiligo	%1.25 (6)	HP:0000010	Recurrent urinary tract infections	%1.25 (6)
HP:0004395	Malnutrition	%1.25 (6)	HP:0003537	Hypouricemia	%1.25 (6)
HP:0033256	Pancolitis	%1.04 (5)	HP:0001933	Subcutaneous hemorrhage	%1.04 (5)
HP:0001890	Autoimmune hemolytic anemia	%1.04 (5)	HP:0003765	Psoriasiform dermatitis	%1.04 (5)
HP:0031813	Colonic eosinophilia	%1.04 (5)	HP:0100507	Reduced blood folate concentration	%1.04 (5)

Tablo 10 (Devamı):

HPO	Tanım	Yüzde (Sayı)	HPO	Tanım	Yüzde (Sayı)
HP:0000238	Hydrocephalus	%1.04 (5)	HP:0002904	Hyperbilirubinemia	%1.04 (5)
HP:0011473	Villous atrophy	%1.04 (5)	HP:0032169	Severe infection	%1.04 (5)
HP:0410151	Eosinophilic infiltration of the esophagus	%1.04 (5)	HP:0100806	Sepsis	%1.04 (5)
HP:0032564	Ileitis	%0.83 (4)	HP:0001596	Alopecia	%0.83 (4)
HP:0003645	Prolonged partial thromboplastin time	%0.83 (4)	HP:0001394	Cirrhosis	%0.83 (4)
HP:0012115	Hepatitis	%0.83 (4)	HP:0001081	Cholelithiasis	%0.83 (4)
HP:0002113	Pulmonary infiltrates	%0.83 (4)	HP:0032016	Abnormal sputum	%0.83 (4)
HP:0012390	Anal fissure	%0.83 (4)	HP:0032551	Hemorrhoids	%0.83 (4)
HP:0032177	Parenchymal consolidation	%0.83 (4)	HP:0040189	Scaling skin	%0.83 (4)
HP:0012393	Allergy	%0.83 (4)	HP:0030148	Heart murmur	%0.83 (4)
HP:0001386	Joint swelling	%0.83 (4)	HP:0030828	Wheezing	%0.83 (4)
HP:0011675	Arrhythmia	%0.83 (4)	HP:0002650	Scoliosis	%0.83 (4)
HP:0032247	Persistent CMV viremia	%0.83 (4)	HP:0000126	Hydronephrosis	%0.83 (4)
HP:0000958	Dry skin	%0.83 (4)	HP:0032174	Tree-in-bud pattern	%0.83 (4)
HP:0004430	Severe combined immunodeficiency	%0.83 (4)	HP:0002091	Restrictive ventilatory defect	%0.83 (4)
HP:0012702	Tenesmus	%0.83 (4)	HP:0002608	Celiac disease	%0.83 (4)
HP:0001025	Urticaria	%0.83 (4)	HP:0001907	Thromboembolism	%0.83 (4)
HP:0012819	Myocarditis	%0.83 (4)	HP:0000819	Diabetes mellitus	%0.83 (4)
HP:0031690	Opportunistic infection	%0.62 (3)	HP:0040075	Hypopituitarism	%0.62 (3)
HP:0000711	Restlessness	%0.62 (3)	HP:0001409	Portal hypertension	%0.62 (3)
HP:0100501	Recurrent bronchiolitis	%0.62 (3)	HP:0025042	Abnormality of mesenteric lymph nodes	%0.62 (3)
HP:0001376	Limitation of joint mobility	%0.62 (3)	HP:0003764	Nevus	%0.62 (3)
HP:0003124	Hypercholesterolemia	%0.62 (3)	HP:0000098	Tall stature	%0.62 (3)
HP:0000028	Cryptorchidism	%0.62 (3)	HP:0002249	Melena	%0.62 (3)
HP:0005266	Intestinal polyp	%0.62 (3)	HP:0030166	Night sweats	%0.62 (3)
HP:0000953	Hyperpigmentation of the skin	%0.62 (3)	HP:0100518	Dysuria	%0.62 (3)
HP:0002722	Recurrent abscess formation	%0.62 (3)	HP:0011956	Intestinal lymphoid nodular hyperplasia	%0.62 (3)
HP:0009763	Limb pain	%0.62 (3)	HP:0011034	Amyloidosis	%0.62 (3)
HP:0000371	Acute otitis media	%0.62 (3)	HP:0002243	Protein-losing enteropathy	%0.62 (3)
HP:0000034	Hydrocele testis	%0.62 (3)	HP:0000938	Osteopenia	%0.62 (3)
HP:0000872	Hashimoto thyroiditis	%0.62 (3)	HP:0100759	Clubbing of fingers	%0.62 (3)
HP:0001631	Atrial septal defect	%0.62 (3)	HP:0002033	Poor suck	%0.62 (3)
HP:0000100	Nephrotic syndrome	%0.62 (3)	HP:0003393	Thenar muscle atrophy	%0.42 (2)
HP:0005528	Bone marrow hypocellularity	%0.42 (2)	HP:0001287	Meningitis	%0.42 (2)
HP:0000510	Rod-cone dystrophy	%0.42 (2)	HP:0002576	Intussusception	%0.42 (2)
HP:0010486	Abnormality of the hypotherar eminence	%0.42 (2)	HP:0032342	Reduced forced expiratory volume in one second	%0.42 (2)

Tablo 10 (Devamı):

HPO	Tanım	Yüzde (Sayı)	HPO	Tanım	Yüzde (Sayı)
HP:0001944	Dehydration	%0.42 (2)	HP:0000252	Microcephaly	%0.42 (2)
HP:0410240	Abnormal circulating IgA level	%0.42 (2)	HP:0002607	Bowel incontinence	%0.42 (2)
HP:0008064	Ichthyosis	%0.42 (2)	HP:0003118	Increased circulating cortisol level	%0.42 (2)
HP:0000405	Conductive hearing impairment	%0.42 (2)	HP:0000790	Hematuria	%0.42 (2)
HP:0000103	Polyuria	%0.42 (2)	HP:0001279	Syncope	%0.42 (2)
HP:0000421	Epistaxis	%0.42 (2)	HP:0006753	Neoplasm of the stomach	%0.42 (2)
HP:0011899	Hyperfibrinogenemia	%0.42 (2)	HP:0002360	Sleep disturbance	%0.42 (2)
HP:0000716	Depression	%0.42 (2)	HP:0002245	Meckel diverticulum	%0.42 (2)
HP:0200043	Verrucae	%0.42 (2)	HP:0032262	Pulmonary tuberculosis	%0.42 (2)
HP:0011370	Recurrent cutaneous fungal infections	%0.42 (2)	HP:0031245	Productive cough	%0.42 (2)
HP:0002570	Steatorrhea	%0.42 (2)	HP:0100646	Thyroiditis	%0.42 (2)
HP:0040090	Abnormality of the tympanic membrane	%0.42 (2)	HP:0012085	Pyuria	%0.42 (2)
HP:0020106	Severe giardiasis	%0.42 (2)	HP:0012117	Hyperalbuminemia	%0.42 (2)
HP:0031020	Bone marrow hypercellularity	%0.42 (2)	HP:0002808	Kyphosis	%0.42 (2)
HP:0004905	Low levels of vitamin A	%0.42 (2)	HP:0001943	Hypoglycemia	%0.42 (2)
HP:0003193	Allergic rhinitis	%0.42 (2)	HP:0031417	Rhinorrhea	%0.42 (2)
HP:0000486	Strabismus	%0.42 (2)	HP:0002665	Lymphoma	%0.42 (2)
HP:0002248	Hematemesis	%0.42 (2)	HP:0002204	Pulmonary embolism	%0.42 (2)
HP:0003281	Increased circulating ferritin concentration	%0.42 (2)	HP:0003259	Elevated circulating creatinine concentration	%0.42 (2)
HP:0002153	Hyperkalemia	%0.42 (2)	HP:0033662	Air bronchogram	%0.42 (2)
HP:0004386	Gastrointestinal inflammation	%0.42 (2)	HP:0001601	Laryngomalacia	%0.42 (2)
HP:0002414	Spina bifida	%0.42 (2)	HP:0001249	Intellectual disability	%0.42 (2)
HP:0025722	Cerebral infarct	%0.42 (2)	HP:0004871	Perineal fistula	%0.42 (2)
HP:0002107	Pneumothorax	%0.42 (2)	HP:0001004	Lymphedema	%0.42 (2)
HP:0010702	Increased circulating antibody level	%0.42 (2)	HP:0011131	Perianal dermatitis	%0.42 (2)
HP:0012343	Decreased circulating ferritin concentration	%0.42 (2)	HP:0002059	Cerebral atrophy	%0.42 (2)
HP:0032271	Extrapulmonary tuberculosis	%0.42 (2)	HP:0000505	Visual impairment	%0.42 (2)
HP:0025388	Thyroid nodule	%0.42 (2)	HP:0034550	Small intestinal lymphoplasmacytic infiltrate	%0.42 (2)
HP:0000952	Jaundice	%0.42 (2)	HP:0100658	Cellulitis	%0.42 (2)
HP:0000135	Hypogonadism	%0.42 (2)	HP:0001271	Polyneuropathy	%0.42 (2)
HP:0410017	Otitis externa	%0.42 (2)	HP:0000999	Pyoderma	%0.21 (1)
HP:0002242	Abnormal intestine morphology	%0.21 (1)	HP:0005197	Generalized morning stiffness	%0.21 (1)
HP:0031700	Invasive parasitic infection	%0.21 (1)	HP:0005588	Patchy palmoplantar hyperkeratosis	%0.21 (1)
HP:0002756	Pathologic fracture	%0.21 (1)	HP:0031207	Hepatic hemangioma	%0.21 (1)

Tablo 10 (Devamı):

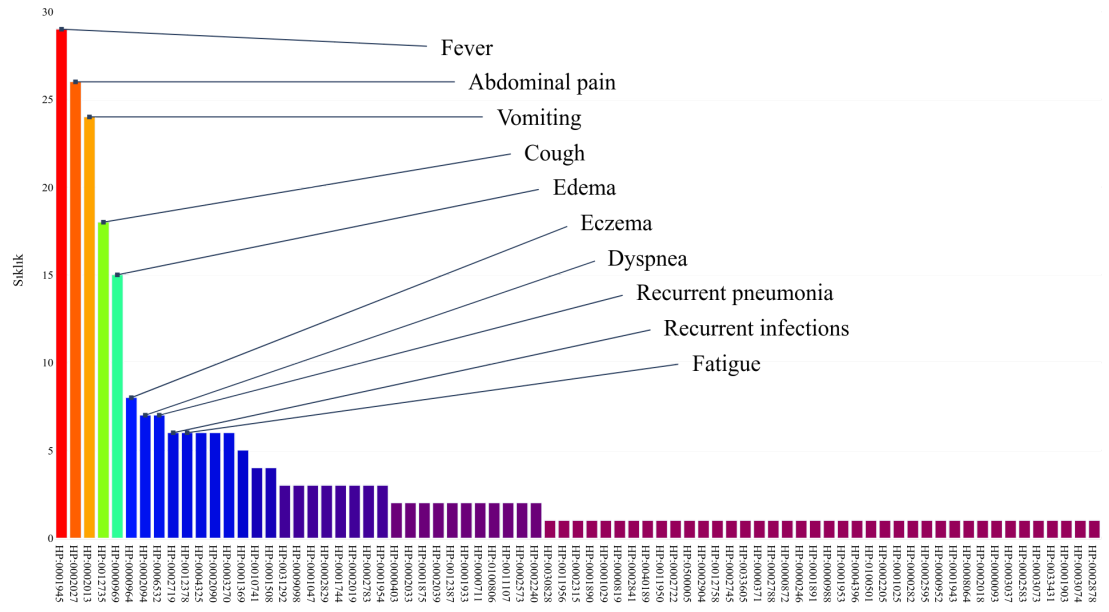
HPO	Tanım	Yüzde (Sayı)	HPO	Tanım	Yüzde (Sayı)
HP:0000739	Anxiety	%0.21 (1)	HP:0010517	Ectopic thymus tissue	%0.21 (1)
HP:0002232	Patchy alopecia	%0.21 (1)	HP:0002516	Increased intracranial pressure	%0.21 (1)
HP:0100581	Dilatation of renal calices	%0.21 (1)	HP:0031379	Abnormal T cell proliferation	%0.21 (1)
HP:0100896	Rectal polyposis	%0.21 (1)	HP:0031970	Abnormal blood urea nitrogen concentration	%0.21 (1)
HP:0000023	Inguinal hernia	%0.21 (1)	HP:0000662	Nyctalopia	%0.21 (1)
HP:0012410	Pure red cell aplasia	%0.21 (1)	HP:0002170	Intracranial hemorrhage	%0.21 (1)
HP:0001803	Nail pits	%0.21 (1)	HP:0100519	Anuria	%0.21 (1)
HP:0100021	Cerebral palsy	%0.21 (1)	HP:0100504	Low levels of vitamin B2	%0.21 (1)
HP:0012538	Gluten intolerance	%0.21 (1)	HP:0011110	Recurrent tonsillitis	%0.21 (1)
HP:0032252	Granuloma	%0.21 (1)	HP:0002837	Recurrent bronchitis	%0.21 (1)
HP:0007305	CNS demyelination	%0.21 (1)	HP:0012731	Ectopic anterior pituitary gland	%0.21 (1)
HP:0000121	Nephrocalcinosis	%0.21 (1)	HP:0000099	Glomerulonephritis	%0.21 (1)
HP:0001776	Bilateral talipes equinovarus	%0.21 (1)	HP:0001085	Papilledema	%0.21 (1)
HP:0001629	Ventricular septal defect	%0.21 (1)	HP:0032066	Decreased serum bicarbonate concentration	%0.21 (1)
HP:0031218	Inappropriate antidiuretic hormone secretion	%0.21 (1)	HP:0031016	Alternating radiolucent and radiodense metaphyseal lines	%0.21 (1)
HP:0031924	Rope sign	%0.21 (1)	HP:0410241	Abnormal circulating IgE level	%0.21 (1)
HP:0500005	Anal pain	%0.21 (1)	HP:0025238	Foot pain	%0.21 (1)
HP:0002605	Hepatic necrosis	%0.21 (1)	HP:0002745	Oral leukoplakia	%0.21 (1)
HP:0034286	Pneumocystis carinii pneumonia	%0.21 (1)	HP:0002863	Myelodysplasia	%0.21 (1)
HP:0002515	Waddling gait	%0.21 (1)	HP:0001733	Pancreatitis	%0.21 (1)
HP:0030991	Sclerosing cholangitis	%0.21 (1)	HP:0003249	Genital ulcers	%0.21 (1)
HP:0000155	Oral ulcer	%0.21 (1)	HP:0033677	Acute respiratory distress syndrome	%0.21 (1)
HP:0001513	Obesity	%0.21 (1)	HP:0410243	Abnormal circulating IgM level	%0.21 (1)
HP:0000047	Hypospadias	%0.21 (1)	HP:0100665	Angioedema	%0.21 (1)
HP:0010976	B lymphocytopenia	%0.21 (1)	HP:0033605	Pustular rash	%0.21 (1)
HP:0002789	Tachypnea	%0.21 (1)	HP:0003929	Ground glass opacity of humeral diaphysis	%0.21 (1)
HP:0100520	Oliguria	%0.21 (1)	HP:0000698	Conical tooth	%0.21 (1)
HP:0000670	Carious teeth	%0.21 (1)	HP:0000138	Ovarian cyst	%0.21 (1)
HP:0000508	Ptosis	%0.21 (1)	HP:0100256	Senile plaques	%0.21 (1)
HP:0100653	Optic neuritis	%0.21 (1)	HP:0000823	Delayed puberty	%0.21 (1)
HP:0003074	Hyperglycemia	%0.21 (1)	HP:0001953	Diabetic ketoacidosis	%0.21 (1)
HP:0005750	Lower-limb joint contracture	%0.21 (1)	HP:0005085	Limited knee flexion/extension	%0.21 (1)
HP:0100628	Esophageal diverticulum	%0.21 (1)	HP:0002036	Hiatus hernia	%0.21 (1)
HP:0033004	Palmar warts	%0.21 (1)	HP:0030812	Enlarged tonsils	%0.21 (1)
HP:0000858	Irregular menstruation	%0.21 (1)	HP:0020072	Persistent EBV viremia	%0.21 (1)
HP:0033650	Pulmonary parenchymal band	%0.21 (1)	HP:0100702	Arachnoid cyst	%0.21 (1)

Tablo 10 (Devamı):

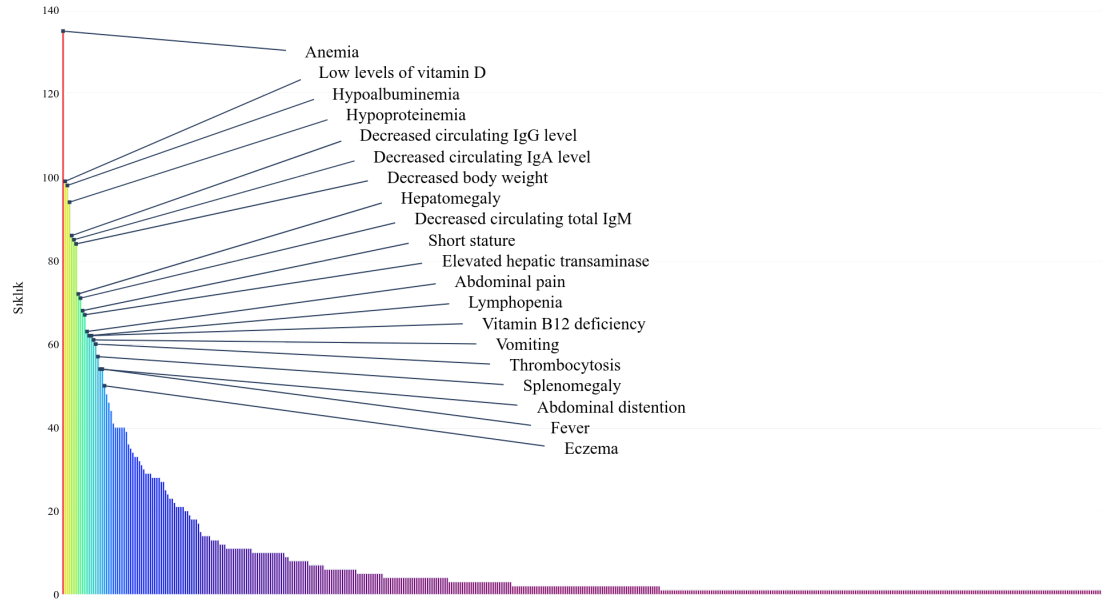
HPO	Tanım	Yüzde (Sayı)	HPO	Tanım	Yüzde (Sayı)
HP:0005202	Helicobacter pylori infection	%0.21 (1)	HP:0002633	Vasculitis	%0.21 (1)
HP:0002206	Pulmonary fibrosis	%0.21 (1)	HP:0007082	Dilated third ventricle	%0.21 (1)
HP:0000089	Renal hypoplasia	%0.21 (1)	HP:0003139	Panhypogammaglobulinemia	%0.21 (1)
HP:0001959	Polydipsia	%0.21 (1)	HP:0001029	Poikiloderma	%0.21 (1)
HP:0003198	Myopathy	%0.21 (1)	HP:0002721	Immunodeficiency	%0.21 (1)
HP:0005193	Restricted large joint movement	%0.21 (1)	HP:0000992	Cutaneous photosensitivity	%0.21 (1)
HP:0001746	Asplenia	%0.21 (1)	HP:0000448	Prominent nose	%0.21 (1)
HP:0011850	Parotitis	%0.21 (1)	HP:0001082	Cholecystitis	%0.21 (1)
HP:0012646	Retractile testis	%0.21 (1)	HP:0000289	Broad philtrum	%0.21 (1)
HP:0012810	Wide nasal base	%0.21 (1)	HP:0010742	Edema of the upper limbs	%0.21 (1)
HP:0033045	Bipedal edema	%0.21 (1)	HP:0000218	High palate	%0.21 (1)
HP:0002848	Decreased specific anti-polysaccharide antibody level	%0.21 (1)	HP:0033537	Mosaic pulmonary attenuation pattern	%0.21 (1)
HP:0000389	Chronic otitis media	%0.21 (1)	HP:0200039	Pustule	%0.21 (1)
HP:0020086	BCGitis	%0.21 (1)	HP:0007209	Facial paralysis	%0.21 (1)
HP:0034269	Abnormal vaginal discharge	%0.21 (1)	HP:0030075	Ductal carcinoma in situ	%0.21 (1)
HP:0012219	Erythema nodosum	%0.21 (1)	HP:0000265	Mastoiditis	%0.21 (1)
HP:0033050	Pharyngalgia	%0.21 (1)	HP:0400007	Polymenorrhea	%0.21 (1)
HP:0100749	Chest pain	%0.21 (1)	HP:0000822	Hypertension	%0.21 (1)
HP:0003401	Paresthesia	%0.21 (1)	HP:0030833	Neck pain	%0.21 (1)
HP:0008249	Thyroid hyperplasia	%0.21 (1)	HP:0100845	Anaphylactic shock	%0.21 (1)
HP:0002726	Recurrent Staphylococcus aureus infections	%0.21 (1)	HP:0000961	Cyanosis	%0.21 (1)
HP:0025538	Palmar edema	%0.21 (1)	HP:0025452	Pyoderma gangrenosum	%0.21 (1)
HP:0000107	Renal cyst	%0.21 (1)	HP:0031141	Increased hepatic echogenicity	%0.21 (1)
HP:0001929	Reduced factor XI activity	%0.21 (1)	HP:0004841	Reduced factor XII activity	%0.21 (1)
HP:0005973	Fructose intolerance	%0.21 (1)	HP:0002563	Constrictive pericarditis	%0.21 (1)
HP:0001562	Oligohydramnios	%0.21 (1)	HP:0002079	Hypoplasia of the corpus callosum	%0.21 (1)
HP:0012848	Small intestinal stenosis	%0.21 (1)	HP:0030168	Dilated superficial abdominal veins	%0.21 (1)
HP:0045040	Abnormal lactate dehydrogenase level	%0.21 (1)	HP:0000492	Abnormal eyelid morphology	%0.21 (1)
HP:0033995	Microvillus inclusions	%0.21 (1)	HP:0000989	Pruritus	%0.21 (1)
HP:0000316	Hypertelorism	%0.21 (1)	HP:0012311	Monocytosis	%0.21 (1)
HP:0031807	Increased basophil count	%0.21 (1)	HP:0005318	Cerebral vasculitis	%0.21 (1)
HP:0002160	Hyperhomocystinemia	%0.21 (1)	HP:0005954	Pulmonary capillary hemangiomatosis	%0.21 (1)
HP:0008326	Reduced circulating vitamin B6 level	%0.21 (1)	HP:0001653	Mitral regurgitation	%0.21 (1)
HP:0034368	Urolithiasis	%0.21 (1)	HP:0000509	Conjunctivitis	%0.21 (1)

Tablo 10 (Devamı):

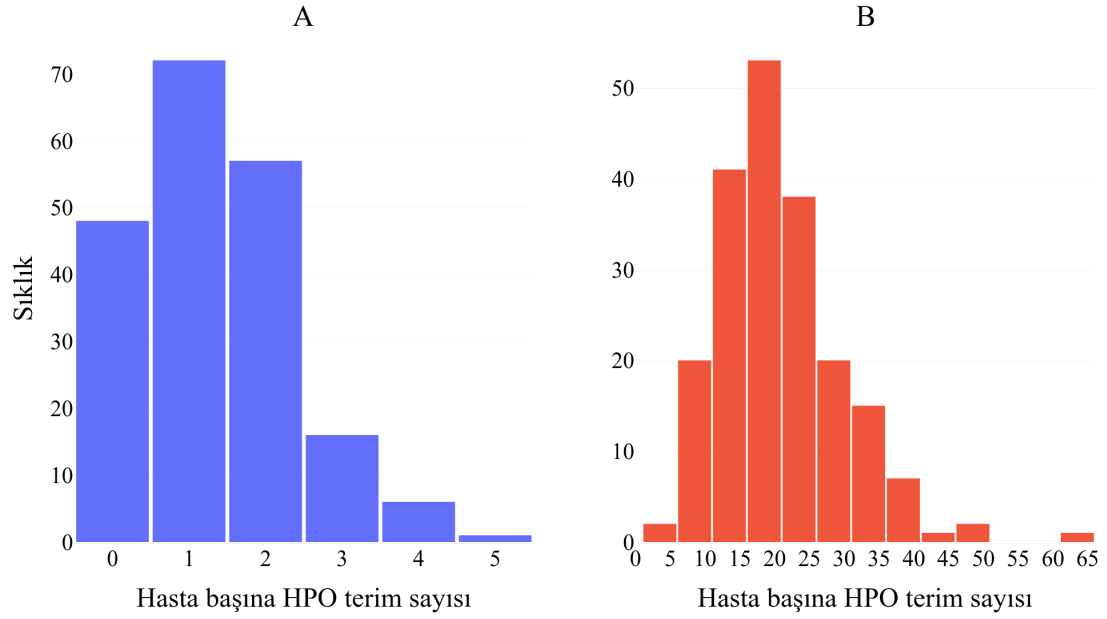
HPO	Tanım	Yüzde (Sayı)	HPO	Tanım	Yüzde (Sayı)
HP:0003236	Elevated circulating creatine kinase concentration	%0.21 (1)	HP:0003037	Enlarged joints	%0.21 (1)
HP:0003146	Hypocholesterolemia	%0.21 (1)	HP:0000518	Cataract	%0.21 (1)
HP:0001396	Cholestasis	%0.21 (1)	HP:0010972	Anemia of inadequate production	%0.21 (1)
HP:0001955	Unexplained fevers	%0.21 (1)	HP:0031693	Severe Epstein Barr virus infection	%0.21 (1)
HP:0001738	Exocrine pancreatic insufficiency	%0.21 (1)	HP:0000554	Uveitis	%0.21 (1)
HP:0000873	Diabetes insipidus	%0.21 (1)	HP:0001094	Iridocyclitis	%0.21 (1)
HP:0032301	Genital warts	%0.21 (1)	HP:0002595	Ileus	%0.21 (1)
HP:0045014	Hypolipidemia	%0.21 (1)	HP:0032203	Lymphoid nodular hyperplasia	%0.21 (1)
HP:0030914	Abnormal peristalsis	%0.21 (1)	HP:0000545	Myopia	%0.21 (1)
HP:0000846	Adrenal insufficiency	%0.21 (1)	HP:0030351	Urticarial plaque	%0.21 (1)
HP:0005425	Recurrent sinopulmonary infections	%0.21 (1)	HP:0010957	Congenital posterior urethral valve	%0.21 (1)
HP:0003326	Myalgia	%0.21 (1)	HP:0001987	Hyperammonemia	%0.21 (1)
HP:0012164	Asterixis	%0.21 (1)	HP:0011896	Subconjunctival hemorrhage	%0.21 (1)
HP:0001698	Pericardial effusion	%0.21 (1)	HP:0002920	Decreased circulating ACTH level	%0.21 (1)
HP:0031414	High serum calcifediol	%0.21 (1)	HP:0001007	Hirsutism	%0.21 (1)
HP:0002293	Alopecia of scalp	%0.21 (1)	HP:0002171	Gliosis	%0.21 (1)
HP:0000622	Blurred vision	%0.21 (1)	HP:0004372	Reduced consciousness/confusion	%0.21 (1)
HP:0001324	Muscle weakness	%0.21 (1)	HP:0000651	Diplopia	%0.21 (1)
HP:0033047	Body ache	%0.21 (1)	HP:0000012	Urinary urgency	%0.21 (1)
HP:0006976	Necrotizing encephalopathy	%0.21 (1)	HP:0001269	Hemiparesis	%0.21 (1)
HP:0001609	Hoarse voice	%0.21 (1)	HP:0012573	Global proximal tubulopathy	%0.21 (1)
HP:0002080	Intention tremor	%0.21 (1)	HP:0410235	Increased anti-insect IgE antibody level	%0.21 (1)
HP:0032310	Granulocytosis	%0.21 (1)	HP:0200114	Metabolic alkalosis	%0.21 (1)
HP:0002015	Dysphagia	%0.21 (1)	HP:0006518	Pulmonary venous occlusion	%0.21 (1)
HP:0034336	Splenic infarction	%0.21 (1)	HP:0002586	Peritonitis	%0.21 (1)
HP:0001028	Hemangioma	%0.21 (1)	HP:0002375	Hypokinesia	%0.21 (1)
HP:0002925	Elevated circulating thyroid-stimulating hormone concentration	%0.21 (1)	HP:0003418	Back pain	%0.21 (1)
HP:0002878	Respiratory failure	%0.21 (1)			



Şekil 29: Hastalık başlangıcındaki semptom ve bulgular
En sık gözlenen 10 HPO teriminin açıklaması verilmiştir.



Şekil 30: Hastalık seyrinde görülen semptom ve bulgular
En sık gözlenen 20 HPO teriminin açıklaması verilmiştir.



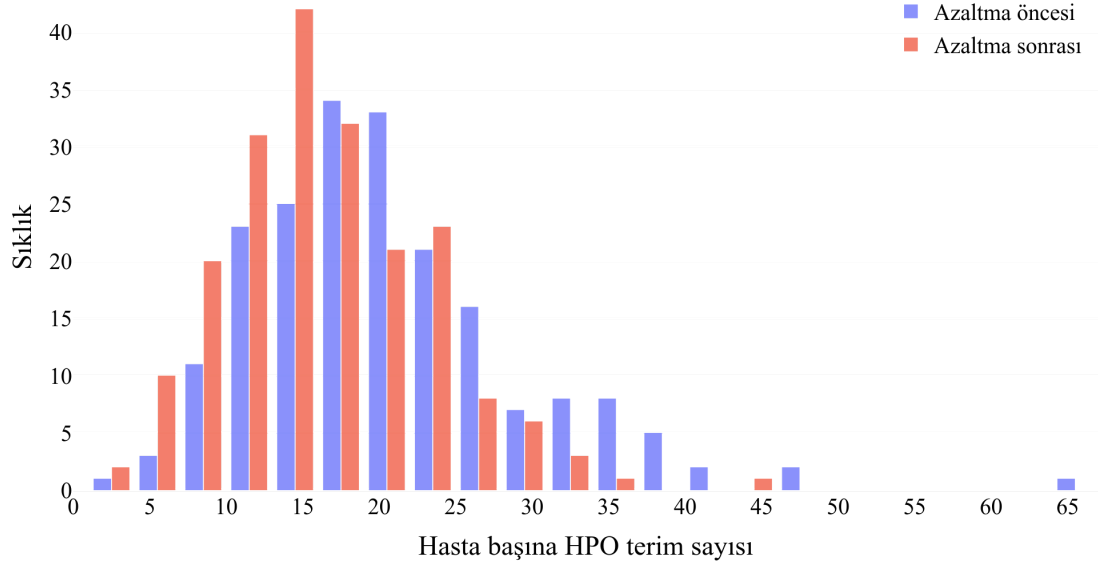
Şekil 31: Hasta başına HPO terimi sayısı

Dahil etme kriterleri olan HPO terimleri çıkarıldıktan sonra. A: Hastalık başlangıcındaki semptom ve bulgular. Dahil etme kriterlerinden sayılan HPO terimleri çıkarıldığında bazı hastalarda 0 başlangıç bulgusu kalmıştır. B: Hastalık seyri boyunca (tanı konulana ve de varsa tedavi başlayana kadar) semptom ve bulgular

Hastaların semptom ve bulgularına dayalı sınıflandırmalarını gerçekleştirmek üzere RF ve GBDT makine öğrenme algoritmaları kullanıldı. Başlangıç semptom ve bulgularına ait HPO terimleri, hastalığın seyri boyunca görülen semptom ve bulguların içinde de olduğu için yanlışlık oluşturulmaması adına makine öğrenimi için kullanılmadı.

Veri setinde bulunan 476 HPO terimine ek olarak cinsiyet, akraba evliliği semptom başlangıç yaşı ve lenfosit alt grupları da eklenerek toplamda 518 farklı değişkenin olduğu bir veri seti oluşturuldu. Değişken sayısı veri setinde tanımlı hasta sayısından fazla olduğu ve çok seyrek ve çok sık gözlenen terimlerin sonuca anlamlı katkıda bulunmayacağı düşünülerek hastaların %5 inden daha azında ve %95'inden daha fazlasında görülen, toplamda 374 HPO terimleri veri setinden çıkarıldı. Kalan 102 HPO teriminin popülasyon içinde görülme sıklığının ortancası 22,5 (IQR 11 -

40); hasta başına tespit edilen HPO terimi sayısı ortancası 15 (IQR 11 - 20) olarak hesaplanmıştır.

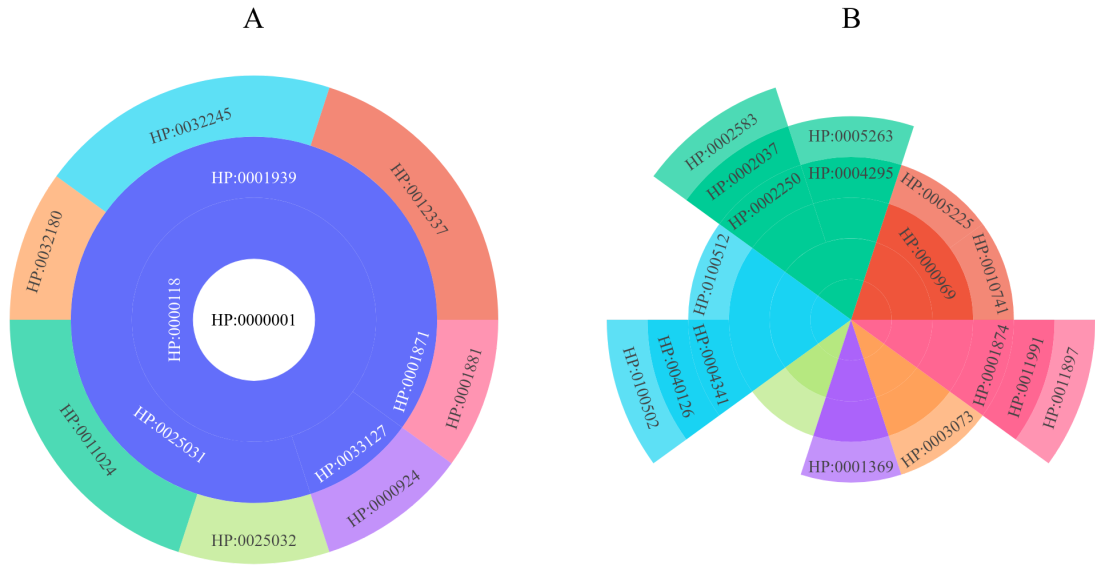


Şekil 32: Terim azaltma öncesi ve sonrası hasta başına tespit edilen HPO terim sayısı dağılımı

Halen fazla sayıda olan değişkenleri daha da azaltmak için HPO birleştirmesi yapılmıştır. Birleştirme süreci, HPO ağacının kök düğümünden başlanarak 4. seviyedeki düğümlerin seçilmesi esasına dayanmaktadır. Her bir HPO teriminin, her aşamada, ilgili ebeveyn düğümlerden birinci sırada olan düğümü takip ederek, birleşeceği seviyedeki atası olan düğüm belirlenmiştir.

HPO birleştirme işlemi sırasında bilgi kaybını en aza indirmek amacıyla, her hastanın her ata düğümüne ilişkin bir ağırlık değeri atanmıştır. Bu ağırlık değerleri hesaplanırken, öncelikle her hastanın sahip olduğu HPO terimleri, bu terimlerin bağlı bulunduğu ata HPO terimleri altında kategorilendirilmiştir. Ardından, her kategoride yer alan HPO terimlerinin sayısı, aynı bireyin toplam HPO terimi sayısına oranlanarak ilgili ata düğümüne yönelik 0 ile 100 arasında bir ağırlık değeri olarak atanmıştır (Şekil 33). Tüm bu işlemler sonucunda HPO terimi sayısı %92,85 azaltılarak 34'e indirilmiştir.

Makine öğrenimi için RF ve GBDT algoritmaları kullanıldı. Her iki algoritma da eksik verileri kabul etmediğinden, eksik veriler değişkenin özelliğine göre dolduruldu. Veri setindeki sürekli değişkenlerdeki eksik veriler için ilgili değişkenin ortalama değeri kullanıldı. Daha sonra, verilerin normalleştirilmesi amacıyla nicel dönüşüm işlemi uygulandı. Kategorik değişkenlerde olan eksik veriler ise, ilgili değişkenin en yaygın gözlenen değeri ile tamamlandı. Lenfosit alt gruplarındaki eksik veriler ise "Normal" olarak işaretlendi.



Şekil 33: HPO terimleri birleştirme süreci

Şekil n119 numaralı hastanın HPO terimlerini birleştirme sürecini göstermektedir.

A: kök düğümünden başlanılarak 4. seviyedeki düğümlerinin belirlenmesini göstermektedir. En dış halkadaki HPO terimlerinin genişliği, o terimin birleştirilme ağırlığı ile orantılıdır. Örneğin HP:0025032 terimi %10 ağırlığa sahipken HP:0011024 terimi %20 ağırlığa sahiptir.

B: Seçilen birleştirme düğümlerinin her birine bağlanan yaprak düğümlerin dağılımını göstermektedir.

Gruplar arasında başarılı ayırım yapabilen bir model bulmak için makine öğrenme algoritmalarının farklı parametreleri, SMOTE ile grup dengelenmesi ve farklı ölçüt yöntemleri seçildi. Toplamda 1320 farklı kombinasyon oluşturuldu. Veri seti küçük olduğu için eğitim ve deneme setleri oluşturulamadı. Onun yerine veri seti

rastgele 3 parçaya bölünerek çapraz doğrulama yöntemi kullanıldı. Seçilen ölçüte göre 0,75 başarılı ayırım yapan modeller başarılı olarak değerlendirildi. Makine öğrenimi modelleri oluşturulmak üzere seçilen değişkenler Tablo 11’de özetlenmiştir.

Algoritmalar, başlangıçta DBK sınıflarını birbirinden ayırmak için test edildi. Ancak, hastaları DBK sınıflarına göre tek bir adımda etkili bir şekilde ayıran bir model belirlemek tespit edilemedi. DBK sınıflarının arasındaki farkları daha kesin bir şekilde anlayabilmek için, her bir DBK sınıfı diğer tüm hastalarla karşılaştırıldı. Bunun yanı sıra, DBK tarafından sınıflandırılan hastalar, DBK dışı mutasyon ve hastalıkların toplamıyla karşılaştırıldı. Ancak, sadece 3 hasta içeren Otoinflamatuvar Hastalıklar sınıfı, kısıtlı hasta sayısı nedeniyle ayrı bir inceleme konusu oluşturmadı.

Tablo 11: Makine öğreniminde kullanılan değişkenler

	RF	GBDT
Kayıp		<i>log loss, exponential</i>
Dallanma kriteri	<i>gini impurity, entropy, log loss</i>	<i>Friedman mean squared error, Squared error</i>
Tahmin edici sayısı	10, 50, 100, 200, 400, 600	
Dal derindiği	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	
Dengeleme	Dengeleme yok, SMOTE	
Ölçüt	f1, ROC-AUC, Jaccard	

Makine öğrenimi ile Sendromik Kombine İmmün Yetmezlikler dışında kalan tüm ayırımlarda başarılı sonuçlar elde edilmiştir. Toplamda 15 farklı model geliştirilmiştir. Bu modellerden 7 tanesi GBDT algoritması, 8 tanesi RF algoritması temellidir. Tüm kombinasyonlarda en iyi sonuç veren ölçüt ROC-AUC’tur. ROC-AUC skorlarının en düşüğü 0,757 ve en yükseği ise 0,929 olduğu görülmüştür.

Tablo 12: Makine öğrenimi modellerinin özellikleri ve sonuçları

Temel algoritma	Karşılaştırılan gruplar	Dengeleme	Skor	Parametreler	Özellikler
GBDT	Hüresel ve Humoral İmmün Yetmezlikler ile diğer tüm hastalar	Yok	0,796	Dallanma Kriteri: Friedman mean squared error Kayıp: Exponential Ağaç derinliği: 2	%13.1 - Abnormal respiratory system morphology %12.37 - Abnormal respiratory system physiology %9.61 - Abnormality of the skin %8.57 - Abnormal ear physiology %7.94 - CD3 ⁺ CD8 ⁺ %
GBDT	Hüresel ve Humoral İmmün Yetmezlikler ile diğer tüm hastalar	SMOTE	0,842	Dallanma Kriteri: Friedman mean squared error Kayıp: Log loss Ağaç derinliği: 2	%20.02 - Abnormal respiratory system morphology %12.43 - Abnormality of digestive system physiology %12.11 - Abnormality of the skin %11.7 - CD4 Efektör Hafıza % %6.13 - Abnormality of the gastrointestinal tract
RFC	Hüresel ve Humoral İmmün Yetmezlikler ile diğer tüm hastalar	Yok	0,859	Dallanma Kriteri: Entropi Ağaç derinliği: 7 Tahmin Edici sayısı: 100	%8.68 - Abnormal respiratory system morphology %8.35 - Abnormality of digestive system physiology %6.96 - Abnormal respiratory system physiology %6.5 - Abnormality of the skin %3.64 - Abnormal abdomen morphology
RFC	Hüresel ve Humoral İmmün Yetmezlikler ile diğer tüm hastalar	SMOTE	0,891	Dallanma Kriteri: Gini impurity Ağaç derinliği: 6 Tahmin edici sayısı: 100	%12.44 - Abnormal respiratory system morphology %8.03 - Abnormality of digestive system physiology %7.33 - Abnormality of the skin %7.18 - Abnormal respiratory system physiology %4.6 - CD4 Efektör Hafıza %
RFC	Antikor Defektleri ile diğer tüm hastalar	Yok	0,771	Dallanma Kriteri: Entropi Ağaç derinliği: 3 Tahmin edici sayısı: 10	%13.24 - Abnormal respiratory system morphology %7.7 - Abnormality of the gastrointestinal tract %6.71 - Abnormality of thrombocytes %6.45 - Abnormal tissue metabolite concentration %5.75 - Abnormal circulating metabolite concentration
RFC	İmmün Düzensizlik Hastalıkları ve diğer tüm hastalar	Yok	0,771	Dallanma Kriteri: Entropi Ağaç derinliği: 5 Tahmin edici sayısı: 100	%8.26 - Naive B % %5.67 - Abnormal homeostasis %4.89 - Abnormal metabolism %3.62 - Abnormality of digestive system physiology %3.24 - Sınıf Değişimi Yapmış Hafıza B %
GBDT	Fagosit Hastalıkları ve diğer tüm hastalar	SMOTE	0,775	Dallanma Kriteri: Friedman mean squared error Kayıp: exponential Ağaç derinliği: 1	%33.77 - Abnormal immune system morphology %18.69 - Abnormal abdomen morphology %14.14 - CD16 ⁺ CD56 ⁺ # %11.39 - Sınıf Değişimi Yapmış Hafıza B % %6.85 - Abnormality of the vasculature
RFC	Fagosit Hastalıkları ve diğer tüm hastalar	SMOTE	0,767	Dallanma Kriteri: Gini impurity Ağaç derinliği: 8 Tahmin edici sayısı: 10	%11.33 - Naive CD8 # %7.77 - Abnormality of the vasculature %7.54 - Abnormal immune system morphology %6.43 - CD8 Efektör Hafıza % %5.25 - Abnormality of the head

Tablo 12 (Devamı):

Temel algoritma	Karşılaştırılan gruplar	Dengeleme	Skor	Parametreler	Özellikler
GBDT	Kompleman Eksiklikleri ve diğer tüm hastalar	Yok	0,827	Dallanma Kriteri: Squared error Kayıp: Log loss Ağaç derinliği: 5	%32.2 - Abnormality of the skeletal system %24.07 - Abnormality of digestive system physiology %12.96 - Growth delay %5.43 - Naive CD4 # %5.05 - Abnormal metabolism
GBDT	Kompleman Eksiklikleri ve diğer tüm hastalar	SMOTE	0,873	Dallanma Kriteri: Friedman mean squared error Kayıp: Log loss Ağaç derinliği: 1	%30.56 - Abnormality of the skeletal system %29.72 - Abnormality of digestive system physiology %17.64 - Abnormal metabolism %4.3 - Abnormal circulating metabolite concentration %4.28 - Abnormal homeostasis
RFC	Kompleman Eksiklikleri ve diğer tüm hastalar	Yok	0,912	Dallanma Kriteri: Entropi Ağaç derinliği: 10 Tahmin edici sayısı: 200	%11.76 - Abnormality of digestive system physiology %10.95 - Abnormal metabolism %6.65 - Abnormal circulating metabolite concentration %6.28 - Abnormality of the skeletal system %4.83 - Abnormal homeostasis
RFC	Kompleman Eksiklikleri ve diğer tüm hastalar	SMOTE	0,930	Dallanma Kriteri: Entropi Ağaç derinliği: 4 Tahmin edici sayısı: 400	%13.45 - Abnormal metabolism %12.26 - Abnormality of digestive system physiology %10.52 - Abnormality of the skeletal system %6.53 - Abnormal homeostasis %6.27 - Abnormal circulating metabolite concentration
GBDT	DBK sınıflamasına dahil olan hastalar ile dahil olmayan hastalar	Yok	0,776	Dallanma Kriteri: Friedman mean squared error Kayıp: Log loss Ağaç derinliği: 6	%24.55 - Abnormality of the gastrointestinal tract %10.18 - Abnormal circulating metabolite concentration %8.56 - Abnormality of coagulation %5.46 - Semptom başlangıç yaşı %4.12 - Abnormality of immune system physiology
GBDT	DBK sınıflamasına dahil olan hastalar ile dahil olmayan hastalar	SMOTE	0,757	Dallanma Kriteri: Friedman mean squared error Kayıp: Log loss Ağaç derinliği: 2	%26.32 - Abnormality of the gastrointestinal tract %11.21 - Abnormal respiratory system morphology %8.56 - CD8 Naive % %7.96 - Akraba evliliği %7.77 - Abnormal circulating metabolite concentration
RFC	DBK sınıflamasına dahil olan hastalar ile dahil olmayan hastalar	Yok	0,780	Dallanma Kriteri: Gini impurity Ağaç derinliği: 8 Tahmin edici sayısı: 50	%14.14 - Abnormality of the gastrointestinal tract %6.29 - Abnormality of immune system physiology %4.7 - Abnormal circulating metabolite concentration %4.45 - Abnormality of body weight %4.1 - Abnormal respiratory system morphology

Özelliklerin yüzdelik değerleri ayırma ne kadar katkı sağladığını göstermektedir.

Tablo 13: Birleştirilmiş HPO terimlerinin farklı DBK sınıflarındaki karşılıkları

DBK Sınıfı	Birleştirilmiş HPO	Yaprak HPO'lar
Hücrel ve Humoral İmmün Yetmezlikler	Abnormal respiratory system morphology	- Pneumonia - Ground-glass opacification - Recurrent pneumonia - Bronchiectasis - Atelectasis - Sinusitis - Pulmonary nodule
	Abnormal respiratory system physiology	- Abnormal breath sound - Asthma - Cough - Dyspnea
	Abnormality of the skin	- Eczema - Atopic dermatitis - Chronic oral candidiasis - Skin rash - Chronic mucocutaneous candidiasis
	Abnormal ear physiology	- Hearing impairment
	Abnormality of digestive system physiology	- Abdominal distention - Abdominal pain - Vomiting - Nausea - Anorexia
	Abnormality of the gastrointestinal tract	- Gastritis - Colitis
	Abnormal abdomen morphology	- Splenomegaly - Hepatomegaly - Ascites
Antikor Defekleri	Abnormal respiratory system morphology	- Pneumonia - Recurrent pneumonia - Bronchiectasis - Sinusitis - Ground-glass opacification - Atelectasis - Pulmonary nodule
	Abnormality of the gastrointestinal tract	- Gastritis - Focal active colitis
	Abnormality of thrombocytes	- Thrombocytopenia - Thrombocytosis
	Abnormal tissue metabolite concentration	- Hepatic steatosis
	Abnormal circulating metabolite concentration	- Hypoproteinemia - Elevated circulating C-reactive protein concentration - Hyperlipidemia - Hypoalbuminemia - Hypomagnesemia - Hypocalcemia

Tablo 13 (Devamı):

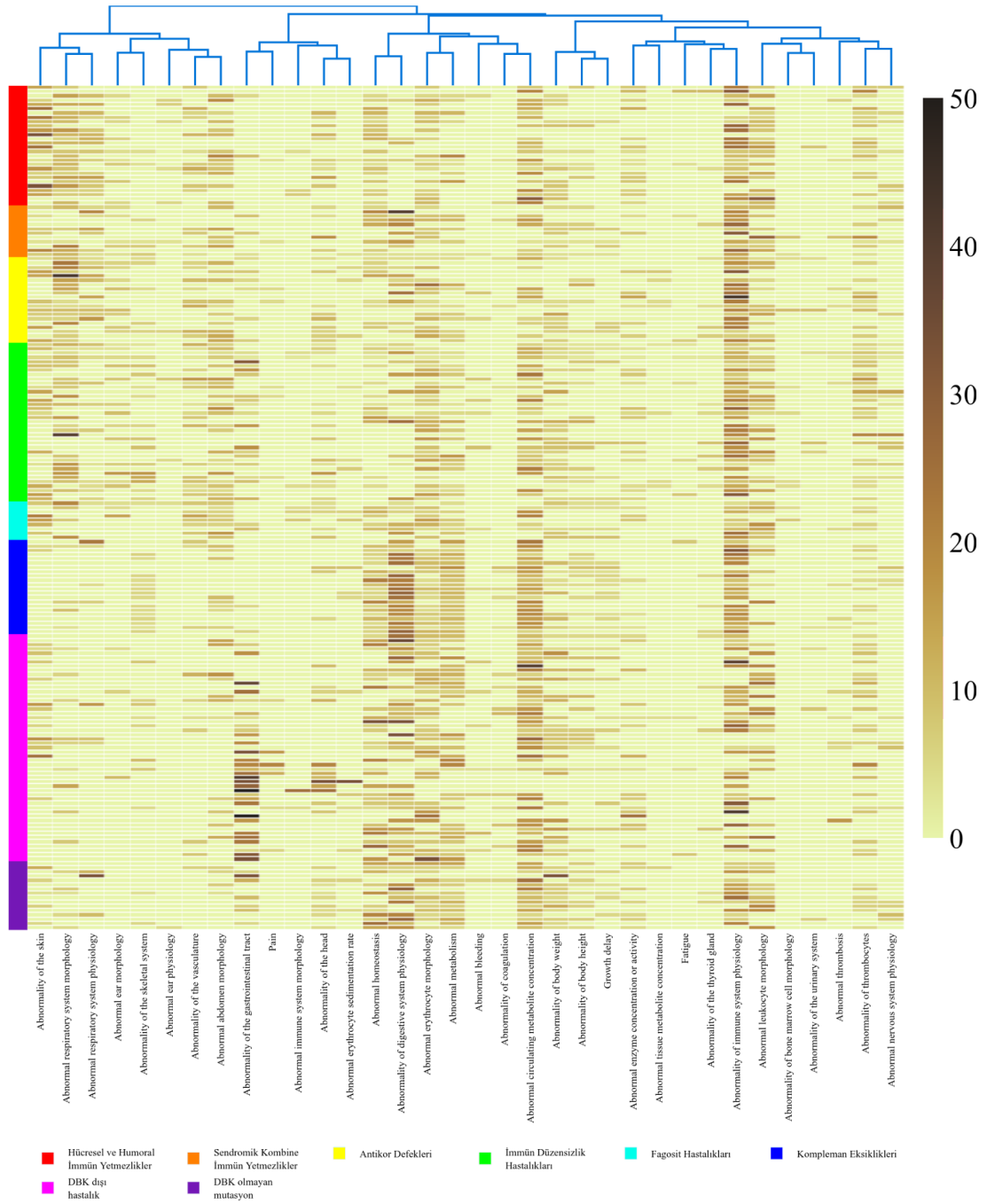
DBK Sınıfı	Birleştirilmiş HPO	Yaprak HPO'lar
İmmün Düzensizlik Hastalıkları	Abnormal homeostasis	- Facial edema - Fever - Edema - Recurrent fever
	Abnormal metabolism	- Low levels of vitamin D - Vitamin B12 deficiency
	Abnormality of digestive system physiology	- Abdominal pain - Nausea - Abdominal distention - Anorexia - Vomiting
Fagosit Hastalıkları	Abnormal immune system morphology	- Abscess
	Abnormal abdomen morphology	- Hepatomegaly - Splenomegaly
	Abnormality of the vasculature	- Lymphadenopathy - Cervical lymphadenopathy - Mediastinal lymphadenopathy - Generalized lymphadenopathy
	Abnormality of the head	- Recurrent upper respiratory tract infections - Recurrent aphthous stomatitis
Kompleman Eksiklikleri	Abnormality of the skeletal system	- Clubbing
	Abnormality of digestive system physiology	- Abdominal distention - Abdominal pain - Anorexia - Nausea - Vomiting
	Growth delay	- Growth delay
	Abnormal metabolism	- Vitamin B12 deficiency - Low levels of vitamin D
	Abnormal circulating metabolite concentration	- Hypoproteinemia - Hyperlipidemia - Hypoalbuminemia - Hypomagnesemia - Hyponatremia - Elevated circulating C-reactive protein concentration - Hypocalcemia
	Abnormal homeostasis	- Facial edema - Edema - Fever - Periorbital edema - Pedal edema - Intestinal edema

Tablo 13 (Devamı):

DBK Sınıfı	Birleştirilmiş HPO	Yaprak HPO'lar
DBK sınıflamasına dahil olan hastalar	Abnormality of the gastrointestinal tract	- Focal active colitis - Ulcerative colitis - Gastritis - Colitis - Gastrointestinal ulcer
	Abnormal circulating metabolite concentration	- Hypoalbuminemia - Hypoproteinemia - Elevated circulating C-reactive protein concentration - Hyperlipidemia - Hypocalcemia - Hyponatremia - Hypomagnesemia
	Abnormality of coagulation	- Prolonged prothrombin time
	Abnormality of immune system physiology	- Decreased circulating antibody level - Food allergy - Decreased circulating IgG level - Increased circulating IgE level - Recurrent infections - Increased circulating IgM level - Decreased circulating IgA level - Decreased circulating total IgM - Crohn's disease - Autoimmunity
	Abnormal respiratory system morphology	- Sinusitis - Atelectasis - Recurrent pneumonia - Pulmonary nodule - Pneumonia
	Abnormality of immune system physiology	- Decreased circulating antibody level - Food allergy - Decreased circulating IgG level - Increased circulating IgE level - Recurrent infections - Increased circulating IgM level - Decreased circulating IgA level - Decreased circulating total IgM - Crohn's disease - Autoimmunity
	Abnormality of body weight	- Decreased body weight - Failure to thrive

Anlatıldığı üzere, makine öğreniminde kullanılan HPO terimleri, çeşitli yaprak düğümlerinin bir araya gelmesiyle oluşturulmuştur. Bu nedenle, belirli bir terim farklı hastalıklarda bazen aynı, bazen de farklı HPO terimlerine karşılık gelebilmektedir. Örneğin “*Abnormality of digestive system physiology*” terimi, Fagosit Hastalıkları sınıfında “*Abdominal pain*” ve “*Vomitting*” terimlerinin birleşiminden oluşurken, Kompleman Eksiklikleri sınıfında “*Abdominal distention*”, “*Abdominal pain*”, “*Anorexia*”, “*Nausea*” ve “*Vomiting*” terimlerinin birleşmesiyle oluşmaktadır. Modeller tarafından seçilen ilk 5 özelliğin detaylı HPO terimleri Tablo 13’te özetlenmiştir.

Seçilen makine öğrenme algoritmalarının temel avantajlarından biri, özellik önemliliklerini göstererek modelin nasıl çalıştığına dair ipuçları sunabilmesidir. Seçilen değişkenlerin hastalık sınıflandırmasındaki rolü, sadece değişkenin hastada mevcut olmasıyla değil, olmamasıyla da ilişkili olabilir. Bu farkı anlamak için değişkenlerin ilgili gruplar içindeki dağılımlarına bakmak da son derece önemlidir. Şekil 34’te, model eğitiminde kullanılan 34 ağırlıklı birleştirilmiş HPO teriminin hastalardaki dağılımları gösterilmiştir. Terimler arasındaki korelasyonlar dendrogramla gösterilmiştir. Hastalar, DBK sınıflarına göre ayrılmıştır.



Şekil 34: HPO birleştirme sonrası ağırlıklı semptom dağılımları
Grafikte satırlar hastaları sütunlar ise birleştirme yapılan HPO terimlerini temsil etmektedir. Kutuların rengi HPO terimlerinin hastadaki ağırlığı (yüzde olarak) arttıkça koyulaşmaktadır.

5. TARTIŞMA

mIBH, erken yaşlarda belirtiler gösteren ve yüksek morbiditeye sahip olan bir hastalık grubunu ifade eder. Temel patolojisi bağırsak inflamasyonu olsa da, hastalarda geniş bir yelpazede fenotipik belirtiler içerebilir. Genetik etyolojinin aydınlatılması ve uygun tedavi seçimi, hastalığın seyrini önemli ölçüde değiştirebilmektedir.

Yapılan istatistiksel analizler sonucunda, çalışmamızda yer alan hastaların boy z skorları, lenfosit sayısı, platelet sayısı, IgG düzeyi, IgA düzeyi ve belli lenfosit alt gruplarının, hastaların ait olduğu DBK ve DBK dışı mIBH tiplerini tahmin etmede yardımcı olduğu belirlenmiştir. Ancak, şu an için tam olarak kesin bir ayırım yapabilen bir yöntem üretilmemiştir. Bu bulgular, hastalık sınıflandırması için potansiyel belirteçlerin tanımlanmasında önemli bir adım sağlamaktadır. Gelecekte bu belirteçler kullanılarak daha fazla veri üzerinde yapılacak analizlerle daha etkili bir ayırım yapabilen yöntemler geliştirmek mümkün olabilir.

Makine öğrenimi yöntemleri kullanılarak gerçekleştirdiğimiz analizler, hastalıkları DBK sınıflarına göre ayırıştırabilen modellerin geliştirilebildiğini göstermiştir. Bu modeller arasında en etkili olanı, özellikle Kompleman Eksiklikleri sınıfını diğer sınıflardan ayırmada üstün performans sergileyen bir modeldir. Ayırım başarısının daha yüksek olmasının sebebi bu sınıfta tek bir genetik hastalık (CHAPLE) bulunması olabilir. Ayrıca, elde edilen sonuçlar, bu modellerin seçtiği ayırım kriterlerinin de dikkat çekici olduğunu ortaya koymaktadır. Bu özelliklerin daha geniş hasta grupları üzerinde incelenmesi, daha genelleyici ve daha geniş bir perspektife sahip olan daha başarılı modellerin oluşturulmasına yardımcı olabilir. Gelecekte daha kapsamlı veri toplama ve analizler ile, hastalık sınıflandırmasında daha yüksek başarı sağlayabilecek modellerin tasarımı ve uygulanması hedeflenebilir.

Görüldüğü üzere erken başlangıçlı bağırsak hastalıklarının klinik özelliklerine bakılarak mutasyonunu tahmin edebilecek güçlü bir yöntem geliştirilememiştir. Bunun temel sebeplerinden biri bunun heterojen ve geniş bir grup olup aslında tek bir hastalık olmamasıdır. Her bir genetik bozukluktan yeterince sayıda hasta bulunması durumunda daha doğru ayırım yapabilecek yöntemlerin keşfedilebileceği düşünülebilir.

Ayrımların başarısız olma sebeplerinden biri de DBK sınıflamasının yapısı olabilir. Bu sınıflama, hastalıkları etkiledikleri immünolojik yolağa göre sınıflandırmaktadır. Ancak, aynı yolağı etkileyen gen mutasyonları her zaman benzer veya yakın fenotipik özellikler sergileyebilir. Bu nedenle, aynı yolağı etkileyen hastalıkları fenotipik özelliklerine göre sınıflandırmak, makine öğrenimi algoritmalarının daha iyi performans göstermesini sağlayabilir. Buna ek olarak, hastalıkların sınıflandırılması süreci de farklı makine öğrenimi algoritmaları kullanılarak daha homojen bir sınıflamaya olanak sağlayabilir. Makine öğrenmesi yöntemi insan faktörünü ortadan kaldırarak daha güvenilir ve tekrarlanabilir sonuçlar elde edilmesini sağlayabilir. Bu noktada ileri düzeyde çalışmalar ve araştırmalar gerekmektedir. Hastalıkların daha etkili ve hassas bir şekilde sınıflandırılması için farklı fenotipik özelliklerin ve genetik verilerin entegrasyonu, makine öğrenme algoritmalarının daha iyi uygulanabilmesi açısından önemli bir adım olabilir.

Panel genetik testler, hızlı sonuç veren ve maliyet etkili genetik analiz yöntemleridir. DBK sınıflamalarına göre düzenlenmiş paneller, klinik uygulamalarda sıkça kullanılmaktadır. Ancak, çalışmanızda da belirtildiği gibi, mİBH şüphesi taşıyan hastalar kolaylıkla DBK sınıflarına dahil edilememektedir. Ayrıca, DBK sınıflamaları içinde yer almayan mİBH vakalarının da varlığı göz önünde bulundurulmalıdır. Bu nedenlerle, mİBH şüphesi olan hastaların panel genetik testlerle taranması, beklenenden daha fazla maliyetli bir süreç haline gelebilmektedir⁵². Özellikle mİBH hastaları için, bilinen mutasyonların tespiti ve yeni mutasyonların tanımlanması amacıyla WES veya WGS gibi daha kapsamlı genetik analiz yöntemlerinin kullanılması önerilmektedir. Bu yaklaşım, daha geniş

bir perspektifle hastalığın genetik temelini anlamamıza ve daha kesin sonuçlar elde etmemize yardımcı olabilir.

mİBH ve PİY, birbirleriyle örtüşen etyoloji ve klinik özelliklere sahip iki hastalık grubunu ifade eder. mİBH'nın tanı ve tedavi sürecine, gastroenteroloji uzmanlarının yanı sıra immünoloji ve genetik uzmanlarının da dahil edilmesi gereklidir. Hastaların değerlendirmesi için hemogram, İg, lenfosit alt grupları ve nötrofil oksidatif patlama çalışması önerilebilir. Özellikle 6 yaş altı hastalar, mutlaka immünolog tarafından değerlendirilmeli; WES / WGS çalışmaları daha erken dönemde planlanmalıdır. Altı yaş üzerinde başlangıç gösteren, dirençli ya da atipik klinik gösteren olgularda da mutlaka genetik çalışma yapılmalıdır ¹⁴⁴.

Çalışılan hasta grubunda, genetik çeşitlilik görülmesine rağmen, hastaların dörtte biri, farklı dört genetik mutasyona sahip vakalardan oluşmaktadır. Bunun temel nedeni, kliniğimize Kompleman Eksikliği (*CD55*), İmmün Disregülasyon (*LRBA*, *CTLA4*) ve Hücrel ve Humoral İmmün Yetmezlikler (*RAG1*) gibi belirli hastalıkların daha sık olarak yönlendirilmesidir. Diğer genetik mutasyonlar ise genellikle daha az sayıda hasta ile temsil edilmiştir. Bu nedenle, hastalıkları değerlendirirken HPO terimlerinde olduğu gibi, hastalıkları tek başlarına değil ait oldukları DBK sınıfına göre sınıflandırmak uygun görülmüştür. Hasta sayılarının ve hastalara ait örneklemelerin az olduğu gruplarda literatürde gösterilen aksine dağılımlar görülebilmektedir. Örneğin Fagosit Hastalıkları grubunda toplamda 9 hasta bulunmaktadır. Bunlardan ikisinin akım sitometrisi verisine ulaşamamıştır. Yine fagositer bozukluk grubunda yer alan G6PC3 eksikliği lenfopeni ile kendini gösteren bir bozukluktur. Kalan hastaların NK hücre oranları normal olsa da hastaların lenfopenik olduğundan, çoğunluğunda sayısal düşüklük varmış gibi görünmektedir.

HPO, insanlar ve makineler tarafından anlaşılabilir bir sınıflandırma sistemidir. Literatürde önerildiği şekilde, genellikle en detaylı terimlerin seçilmesi tavsiye edilmektedir. Ancak çalışmanızda çok sayıda HPO terimi ve daha az hasta örneği bulunması, birleştirme tekniklerini tercih etmenize yol açmıştır. HPO, yaklaşık her 2

ayda bir güncellenmektedir. Bu güncellemeler sırasında yeni terimler eklenmekte, mevcut terimler çıkarılmakta veya birleştirilmekte, ontoloji ağacının yapısı değişmekte ve dallanmalar oluşturulabilmektedir. Bu nedenle, birleştirme yöntemleri sonucunda elde edilen terim ağırlıkları ve terimlerin, farklı HPO sürümlerinde aynı olmayabileceği göz önünde bulundurulmalıdır. HPO ağacının sürekli olarak değişiyor olması ve aynı şekilde DBK sınıflamasının da gelecekte değişebileceği gerçeği, üretilen modellerin pratik uygulamada kullanımına uygunluğunu sınırlayabilir. Bu durum, model sonuçlarını klinik kararlar için güvenilir bir şekilde kullanmak isteyen uzmanlar için önemli bir husustur. Bu noktada, model güvenilirliğini ve sürekli kullanılabilirliğini sağlamak için yöntemlerinizi geliştirmek ve güncellemek önemli olacaktır.

Araştırma evrenimizde çoğu genetik bozukluk yalnız 1 hasta ile temsil edilmiştir. Bu nedenle de analizler her bir genin farklılıklarını tespit etmekten ziyade genin ait olduğu DBK sınıfını, etkilenen sistemi, tahmin etmek üzerine yapılmıştır. Her ne kadar aynı sistem etkilense de aynı sınıfa dahil olan hastalıkların fenotipik özelliklerinin birbiri ile benzer olması gerekliliği yoktur. Bu nedenle birbiriyle zıt iki HPO terimi aynı sınıfa dahil hastalarda görülebilmektedir. Bunu sonucu olarak da tez içinde kullandığımız HPO birleştirme işlemi sırasında bilgi kaybı olabilmektedir. Bu kayıp modellerin başarısını etkileyebilir.

Çalışmamızda bir diğer önemli eksiklik hastaların doku tanıılarının çalışmaya eklenmemiş olmasıdır. Hastanemiz Çocuk Gastroenteroloji bölümünde yapılan bir çalışmada da gösterildiği üzere PİY olup ağır gastrointestinal semptom gösteren hastaların doku incelemelerinde İBH ile uyumlu sonuçlar gözlenmiştir ¹⁴⁵. Doku tanıılarının eksik olması makine öğrenim modellerin tahmin başarısını düşürdüğü düşünülmektedir.

6. KAYNAKÇA

1. Bolton C, Smillie CS, Pandey S, et al. An integrated taxonomy for monogenic inflammatory bowel disease. *Gastroenterology*. 2022;162(3):859-876. doi:10.1053/j.gastro.2021.11.014
2. Girardelli M, Basaldella F, Paolera SD, et al. Genetic profile of patients with early onset inflammatory bowel disease. *Gene*. 2018;645:18-29. doi:10.1016/j.gene.2017.12.029
3. Pazmandi J, Kalinichenko A, Ardy RC, Boztug K. Early-onset inflammatory bowel disease as a model disease to identify key regulators of immune homeostasis mechanisms. *Immunol Rev*. 2019;287(1):162-185. doi:10.1111/imr.12726
4. Tangye SG, Al-Herz W, Bousfiha A, et al. Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. *J Clin Immunol*. 2022;42(7):1473-1507. doi:10.1007/s10875-022-01289-3
5. Ye Z, Huang Y, Wang Y, Lu J, Wu J, Yu Z. Phenotype and Genotype of a Cohort of Chinese Children with Early-Onset Protein-Losing Enteropathy. *J Pediatr*. 2019;208:38-42.e3. doi:10.1016/j.jpeds.2018.12.003
6. Ozen A, Comrie WA, Ardy RC, et al. CD55 Deficiency, Early-Onset Protein-Losing Enteropathy, and Thrombosis. *N Engl J Med*. 2017;377(1):52-61. doi:10.1056/NEJMoa1615887
7. Zhang ZZ, Zhang Y, He T, et al. Homozygous IL37 mutation associated with infantile inflammatory bowel disease. *Proc Natl Acad Sci USA*. 2021;118(10). doi:10.1073/pnas.2009217118
8. Lucas CL, Zhang Y, Venida A, et al. Heterozygous splice mutation in PIK3R1 causes human immunodeficiency with lymphoproliferation due to dominant activation of PI3K. *J Exp Med*. 2014;211(13):2537-2547. doi:10.1084/jem.20141759
9. Yao Y, Kim G, Shafer S, et al. Mucus sialylation determines intestinal host-commensal homeostasis. *Cell*. 2022;185(7):1172-1188.e28. doi:10.1016/j.cell.2022.02.013
10. Kubo S, Fritz JM, Raquer-McKay HM, et al. Congenital iRHOM2 deficiency causes ADAM17 dysfunction and environmentally directed immunodysregulatory disease. *Nat Immunol*. 2022;23(1):75-85.

doi:10.1038/s41590-021-01093-y

11. Köhler S, Gargano M, Matentzoglou N, et al. The human phenotype ontology in 2021. *Nucleic Acids Res.* 2021;49(D1):D1207-D1217. doi:10.1093/nar/gkaa1043
12. Levitt DG, Levitt MD. Protein losing enteropathy: comprehensive review of the mechanistic association with clinical and subclinical disease states. *Clin Exp Gastroenterol.* 2017;10:147-168. doi:10.2147/CEG.S136803
13. Waldmann TA, Steinfeld JL, Dutcher TF, Davidson JD, Gordon RS. The role of the gastrointestinal system in “idiopathic hypoproteinemia”. *Gastroenterology.* 1961;41:197-207. doi:10.1016/S0016-5085(19)35130-3
14. Braamskamp MJAM, Dolman KM, Tabbers MM. Clinical practice. Protein-losing enteropathy in children. *Eur J Pediatr.* 2010;169(10):1179-1185. doi:10.1007/s00431-010-1235-2
15. Bennish ML, Salam MA, Wahed MA. Enteric protein loss during shigellosis. *Am J Gastroenterol.* 1993;88(1):53-57.
16. Sutton DL, Kamath KR. Giardiasis with protein-losing enteropathy. *J Pediatr Gastroenterol Nutr.* 1985;4(1):56-59. doi:10.1097/00005176-198502000-00011
17. Goos G, Declercq P, Proesmans W. Serum proteins in the haemolytic uraemic syndrome. *Pediatr Nephrol.* 1995;9(3):292-294. doi:10.1007/BF02254186
18. Chehade M, Magid MS, Mofidi S, Nowak-Węgrzyn A, Sampson HA, Sicherer SH. Allergic eosinophilic gastroenteritis with protein-losing enteropathy: intestinal pathology, clinical course, and long-term follow-up. *J Pediatr Gastroenterol Nutr.* 2006;42(5):516-521. doi:10.1097/01.mpg.0000221903.61157.4e
19. Maloney J, Nowak-Węgrzyn A. Educational clinical case series for pediatric allergy and immunology: allergic proctocolitis, food protein-induced enterocolitis syndrome and allergic eosinophilic gastroenteritis with protein-losing gastroenteropathy as manifestations of non-IgE-mediated cow’s milk allergy. *Pediatr Allergy Immunol.* 2007;18(4):360-367. doi:10.1111/j.1399-3038.2007.00561.x
20. Ingle SB, Hinge Ingle CR. Primary intestinal lymphangiectasia: Minireview. *World J Clin Cases.* 2014;2(10):528-533. doi:10.12998/wjcc.v2.i10.528
21. Vignes S, Bellanger J. Primary intestinal lymphangiectasia (Waldmann’s disease). *Orphanet J Rare Dis.* 2008;3:5. doi:10.1186/1750-1172-3-5
22. Fuss IJ, Strober W, Cuccherini BA, et al. Intestinal lymphangiectasia, a disease characterized by selective loss of naive CD45RA+ lymphocytes into the gastrointestinal tract. *Eur J Immunol.* 1998;28(12):4275-4285.

doi:10.1002/(SICI)1521-4141(199812)28:12<4275::AID-IMMU4275>3.0.CO;2-P

23. Sanges S, Germain N, Vignes S, et al. Protein-losing Enteropathy as a Complication and/or Differential Diagnosis of Common Variable Immunodeficiency. *J Clin Immunol*. 2022;42(7):1461-1472. doi:10.1007/s10875-022-01299-1
24. Rubalcava NS, Gadepalli SK. Inflammatory bowel disease in children and adolescents. *Adv Pediatr*. 2021;68:121-142. doi:10.1016/j.yapd.2021.05.005
25. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature*. 2011;474(7351):307-317. doi:10.1038/nature10209
26. Burisch J, Jess T, Martinato M, Lakatos PL, ECCO -EpiCom. The burden of inflammatory bowel disease in Europe. *J Crohns Colitis*. 2013;7(4):322-337. doi:10.1016/j.crohns.2013.01.010
27. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet*. 2017;390(10114):2769-2778. doi:10.1016/S0140-6736(17)32448-0
28. Johnston RD, Logan RFA. What is the peak age for onset of IBD? *Inflamm Bowel Dis*. 2008;14 Suppl 2:S4-5. doi:10.1002/ibd.20545
29. Baldassano RN, Piccoli DA. Inflammatory bowel disease in pediatric and adolescent patients. *Gastroenterol Clin North Am*. 1999;28(2):445-458. doi:10.1016/s0889-8553(05)70064-9
30. Heyman MB, Kirschner BS, Gold BD, et al. Children with early-onset inflammatory bowel disease (IBD): analysis of a pediatric IBD consortium registry. *J Pediatr*. 2005;146(1):35-40. doi:10.1016/j.jpeds.2004.08.043
31. Yu YR, Rodriguez JR. Clinical presentation of Crohn's, ulcerative colitis, and indeterminate colitis: Symptoms, extraintestinal manifestations, and disease phenotypes. *Semin Pediatr Surg*. 2017;26(6):349-355. doi:10.1053/j.sempedsurg.2017.10.003
32. UpToDate - Clinical presentation and diagnosis of inflammatory bowel disease in children. Accessed April 26, 2023. <https://www.uptodate.com/contents/clinical-presentation-and-diagnosis-of-inflammatory-bowel-disease-in-children>
33. Greuter T, Bertoldo F, Rechner R, et al. Extraintestinal Manifestations of Pediatric Inflammatory Bowel Disease: Prevalence, Presentation, and Anti-TNF Treatment. *J Pediatr Gastroenterol Nutr*. 2017;65(2):200-206. doi:10.1097/MPG.0000000000001455
34. Veloso FT, Carvalho J, Magro F. Immune-related systemic manifestations of

- inflammatory bowel disease. A prospective study of 792 patients. *J Clin Gastroenterol*. 1996;23(1):29-34. doi:10.1097/00004836-199607000-00009
35. Kelsen JR, Sullivan KE. Inflammatory bowel disease in primary immunodeficiencies. *Curr Allergy Asthma Rep*. 2017;17(8):57. doi:10.1007/s11882-017-0724-z
 36. Levine A. Pediatric inflammatory bowel disease: is it different? *Dig Dis*. 2009;27(3):212-214. doi:10.1159/000228552
 37. Glickman JN, Bousvaros A, Farraye FA, et al. Pediatric patients with untreated ulcerative colitis may present initially with unusual morphologic findings. *Am J Surg Pathol*. 2004;28(2):190-197. doi:10.1097/00000478-200402000-00006
 38. Galbraith SS, Drolet BA, Kugathasan S, Paller AS, Esterly NB. Asymptomatic inflammatory bowel disease presenting with mucocutaneous findings. *Pediatrics*. 2005;116(3):e439-44. doi:10.1542/peds.2004-2281
 39. Müller KE, Lakatos PL, Arató A, et al. Incidence, Paris classification, and follow-up in a nationwide incident cohort of pediatric patients with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr*. 2013;57(5):576-582. doi:10.1097/MPG.0b013e31829f7d8c
 40. UpToDate - Overview of the management of Crohn disease in children and adolescents. Overview of the management of Crohn disease in children and adolescents. Accessed May 19, 2023. <https://www.uptodate.com/contents/overview-of-the-management-of-crohn-disease-in-children-and-adolescents>
 41. Levine A, Kugathasan S, Annese V, et al. Pediatric onset Crohn's colitis is characterized by genotype-dependent age-related susceptibility. *Inflamm Bowel Dis*. 2007;13(12):1509-1515. doi:10.1002/ibd.20244
 42. Kelsen JR, Russo P, Sullivan KE. Early-Onset Inflammatory Bowel Disease. *Immunol Allergy Clin North Am*. 2019;39(1):63-79. doi:10.1016/j.iac.2018.08.008
 43. Yang H, McElree C, Roth MP, Shanahan F, Targan SR, Rotter JI. Familial empirical risks for inflammatory bowel disease: differences between Jews and non-Jews. *Gut*. 1993;34(4):517-524. doi:10.1136/gut.34.4.517
 44. Nambu R, Warner N, Mulder DJ, et al. A systematic review of monogenic inflammatory bowel disease. *Clin Gastroenterol Hepatol*. 2022;20(4):e653-e663. doi:10.1016/j.cgh.2021.03.021
 45. Levine A, Griffiths A, Markowitz J, et al. Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. *Inflamm Bowel Dis*. 2011;17(6):1314-1321.

doi:10.1002/ibd.21493

46. de Bie CI, Buderus S, Sandhu BK, et al. Diagnostic workup of paediatric patients with inflammatory bowel disease in Europe: results of a 5-year audit of the EUROKIDS registry. *J Pediatr Gastroenterol Nutr.* 2012;54(3):374-380. doi:10.1097/MPG.0b013e318231d984
47. Cannioto Z, Berti I, Martelossi S, et al. IBD and IBD mimicking enterocolitis in children younger than 2 years of age. *Eur J Pediatr.* 2009;168(2):149-155. doi:10.1007/s00431-008-0721-2
48. Levine A, Koletzko S, Turner D, et al. ESPGHAN revised porto criteria for the diagnosis of inflammatory bowel disease in children and adolescents. *J Pediatr Gastroenterol Nutr.* 2014;58(6):795-806. doi:10.1097/MPG.0000000000000239
49. Muise AM, Snapper SB, Kugathasan S. The age of gene discovery in very early onset inflammatory bowel disease. *Gastroenterology.* 2012;143(2):285-288. doi:10.1053/j.gastro.2012.06.025
50. Uhlig HH, Schwerdt T, Koletzko S, et al. The diagnostic approach to monogenic very early onset inflammatory bowel disease. *Gastroenterology.* 2014;147(5):990-1007.e3. doi:10.1053/j.gastro.2014.07.023
51. Kanegane H. Inflammatory bowel diseases and primary immunodeficiency diseases. *Immunol Med.* 2018;41(4):154-161. doi:10.1080/25785826.2018.1556025
52. Uhlig HH, Charbit-Henrion F, Kotlarz D, et al. Clinical genomics for the diagnosis of monogenic forms of inflammatory bowel disease: A position paper from the paediatric IBD porto group of european society of paediatric gastroenterology, hepatology and nutrition. *J Pediatr Gastroenterol Nutr.* 2021;72(3):456-473. doi:10.1097/MPG.0000000000003017
53. Notarangelo LD, Gambineri E, Badolato R. Immunodeficiencies with autoimmune consequences. *Adv Immunol.* 2006;89:321-370. doi:10.1016/S0065-2776(05)89008-X
54. Giardino G, Gallo V, Prencipe R, et al. Unbalanced immune system: immunodeficiencies and autoimmunity. *Front Pediatr.* 2016;4:107. doi:10.3389/fped.2016.00107
55. Amaya-Uribe L, Rojas M, Azizi G, Anaya J-M, Gershwin ME. Primary immunodeficiency and autoimmunity: A comprehensive review. *J Autoimmun.* 2019;99:52-72. doi:10.1016/j.jaut.2019.01.011
56. Bonilla FA, Khan DA, Ballas ZK, et al. Practice parameter for the diagnosis and management of primary immunodeficiency. *J Allergy Clin Immunol.* 2015;136(5):1186-205.e1. doi:10.1016/j.jaci.2015.04.049

57. Milner JD, Holland SM. The cup runneth over: lessons from the ever-expanding pool of primary immunodeficiency diseases. *Nat Rev Immunol.* 2013;13(9):635-648. doi:10.1038/nri3493
58. 10 Warning Signs of Primary Immunodeficiency. 10 Warning Signs of Primary Immunodeficiency. Accessed June 14, 2023. <https://info4pi.org/library/educational-materials/>
59. Devonshire AL, Makhija M. Approach to primary immunodeficiency. *Allergy Asthma Proc.* 2019;40(6):465-469. doi:10.2500/aap.2019.40.4273
60. UpToDate - IUIS Classification Summary. Accessed May 24, 2023. <https://www.uptodate.com/contents/inborn-errors-of-immunity-primary-immunodeficiencies-classification>
61. UpToDate - Immune and microbial mechanisms in the pathogenesis of inflammatory bowel disease. Accessed May 24, 2023. <https://www.uptodate.com/contents/immune-and-microbial-mechanisms-in-the-pathogenesis-of-inflammatory-bowel-disease>
62. Mayer L, Shlien R. Evidence for function of Ia molecules on gut epithelial cells in man. *J Exp Med.* 1987;166(5):1471-1483. doi:10.1084/jem.166.5.1471
63. Mayer L, Eisenhardt D, Salomon P, Bauer W, Plous R, Piccinini L. Expression of class II molecules on intestinal epithelial cells in humans. Differences between normal and inflammatory bowel disease. *Gastroenterology.* 1991;100(1):3-12. doi:10.1016/0016-5085(91)90575-6
64. Kamada N, Hisamatsu T, Okamoto S, et al. Unique CD14 intestinal macrophages contribute to the pathogenesis of Crohn disease via IL-23/IFN-gamma axis. *J Clin Invest.* 2008;118(6):2269-2280. doi:10.1172/JCI34610
65. Reimund JM, Wittersheim C, Dumont S, et al. Increased production of tumour necrosis factor-alpha interleukin-1 beta, and interleukin-6 by morphologically normal intestinal biopsies from patients with Crohn's disease. *Gut.* 1996;39(5):684-689. doi:10.1136/gut.39.5.684
66. Takayama T, Kamada N, Chinen H, et al. Imbalance of NKp44(+)NKp46(-) and NKp44(-)NKp46(+) natural killer cells in the intestinal mucosa of patients with Crohn's disease. *Gastroenterology.* 2010;139(3):882-892, 892.e1. doi:10.1053/j.gastro.2010.05.040
67. Mitsialis V, Wall S, Liu P, et al. Single-Cell Analyses of Colon and Blood Reveal Distinct Immune Cell Signatures of Ulcerative Colitis and Crohn's Disease. *Gastroenterology.* 2020;159(2):591-608.e10. doi:10.1053/j.gastro.2020.04.074

68. Fiocchi C, Battisto JR, Farmer RG. Studies on isolated gut mucosal lymphocytes in inflammatory bowel disease. Detection of activated T cells and enhanced proliferation to Staphylococcus aureus and lipopolysaccharides. *Dig Dis Sci*. 1981;26(8):728-736. doi:10.1007/BF01316863
69. Panda SK, Colonna M. Innate lymphoid cells in mucosal immunity. *Front Immunol*. 2019;10:861. doi:10.3389/fimmu.2019.00861
70. Camus M, Esses S, Pariente B, et al. Oligoclonal expansions of mucosal T cells in Crohn's disease predominate in NKG2D-expressing CD4 T cells. *Mucosal Immunol*. 2014;7(2):325-334. doi:10.1038/mi.2013.51
71. Chott A, Probert CS, Gross GG, Blumberg RS, Balk SP. A common TCR beta-chain expressed by CD8+ intestinal mucosa T cells in ulcerative colitis. *J Immunol*. 1996;156(8):3024-3035.
72. Uo M, Hisamatsu T, Miyoshi J, et al. Mucosal CXCR4+ IgG plasma cells contribute to the pathogenesis of human ulcerative colitis through FcγR-mediated CD14 macrophage activation. *Gut*. 2013;62(12):1734-1744. doi:10.1136/gutjnl-2012-303063
73. Fuss IJ, Marth T, Neurath MF, Pearlstein GR, Jain A, Strober W. Anti-interleukin 12 treatment regulates apoptosis of Th1 T cells in experimental colitis in mice. *Gastroenterology*. 1999;117(5):1078-1088. doi:10.1016/s0016-5085(99)70392-6
74. Powrie F, Leach MW, Mauze S, Menon S, Caddle LB, Coffman RL. Inhibition of Th1 responses prevents inflammatory bowel disease in scid mice reconstituted with CD45RBhi CD4+ T cells. *Immunity*. 1994;1(7):553-562. doi:10.1016/1074-7613(94)90045-0
75. Targan SR, Hanauer SB, van Deventer SJ, et al. A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group. *N Engl J Med*. 1997;337(15):1029-1035. doi:10.1056/NEJM199710093371502
76. Heller F, Fuss IJ, Nieuwenhuis EE, Blumberg RS, Strober W. Oxazolone colitis, a Th2 colitis model resembling ulcerative colitis, is mediated by IL-13-producing NK-T cells. *Immunity*. 2002;17(5):629-638. doi:10.1016/s1074-7613(02)00453-3
77. Iqbal N, Oliver JR, Wagner FH, Lazenby AS, Elson CO, Weaver CT. T helper 1 and T helper 2 cells are pathogenic in an antigen-specific model of colitis. *J Exp Med*. 2002;195(1):71-84. doi:10.1084/jem.2001889
78. Strober W, Fuss IJ. Proinflammatory cytokines in the pathogenesis of inflammatory bowel diseases. *Gastroenterology*. 2011;140(6):1756-1767. doi:10.1053/j.gastro.2011.02.016

79. Picard C, Bobby Gaspar H, Al-Herz W, et al. International union of immunological societies: 2017 primary immunodeficiency diseases committee report on inborn errors of immunity. *J Clin Immunol*. 2018;38(1):96-128. doi:10.1007/s10875-017-0464-9
80. Tegtmeier D, Seidl M, Gerner P, Baumann U, Klemann C. Inflammatory bowel disease caused by primary immunodeficiencies-Clinical presentations, review of literature, and proposal of a rational diagnostic algorithm. *Pediatr Allergy Immunol*. 2017;28(5):412-429. doi:10.1111/pai.12734
81. ontology | Etymology, origin and meaning of ontology by etymonline. Accessed April 26, 2023. https://www.etymonline.com/word/ontology#etymonline_v_7042
82. Gruber TR. Toward principles for the design of ontologies used for knowledge sharing? *Int J Hum Comput Stud*. 1995;43(5-6):907-928. doi:10.1006/ijhc.1995.1081
83. Ashburner M, Ball CA, Blake JA, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*. 2000;25(1):25-29. doi:10.1038/75556
84. OBO Foundry. Accessed April 26, 2023. <http://obofoundry.org/>
85. Online Mendelian Inheritance in Man. Online Mendelian Inheritance in Man. Accessed April 26, 2023. <https://www.omim.org/>
86. Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet*. 2008;83(5):610-615. doi:10.1016/j.ajhg.2008.09.017
87. Köhler S, Øien NC, Buske OJ, et al. Encoding Clinical Data with the Human Phenotype Ontology for Computational Differential Diagnostics. *Curr Protoc Hum Genet*. 2019;103(1):e92. doi:10.1002/cphg.92
88. Kernohan KD, Hartley T, Alirezaie N, et al. Evaluation of exome filtering techniques for the analysis of clinically relevant genes. *Hum Mutat*. 2018;39(2):197-201. doi:10.1002/humu.23374
89. Exomiser documentation. Exomiser documentation. Accessed April 21, 2023. <https://exomiser.readthedocs.io/en/latest/>
90. The Phenomizer - Clinical Diagnostics with Similarity Searches in Ontologies. The Phenomizer - Clinical Diagnostics with Similarity Searches in Ontologies. Accessed April 26, 2023. <https://compbio.charite.de/phenomizer/>
91. Orphanet. Orphanet. Accessed April 27, 2023. <https://clinicalsigns.orphanet.app/>

92. Shapiro SC. *Encyclopedia of Artificial Intelligence*. 2 Volume Set. Wiley; 1992:1792.
93. Clark W, Golinski Jan, Schaffer S. *The Sciences in Enlightened Europe*. 1st ed. University of Chicago Press; 1999:580.
94. Legendre AM. *Nouvelles Méthodes Pour La Détermination Des Orbites Des Comètes (French Edition)*. Wentworth Press; 2018:288.
95. Turing AM. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*. 1937;s2-42(1):230-265. doi:10.1112/plms/s2-42.1.230
96. Mitchell TM. *Machine Learning*. 1st ed. McGraw-Hill Education; 1997:432.
97. Allen-b-tucker. *Computer-Science-and-Engineering-Handbook*. 2nd ed. Taylor & Francis; 2004:2752.
98. Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: A survey. *jair*. 1996;4:237-285. doi:10.1613/jair.301
99. Studer M, Ritschard G, Gabadinho A, Müller NS. Discrepancy analysis of state sequences. *Sociol Methods Res*. 2011;40(3):471-510. doi:10.1177/0049124111415372
100. Wu X, Kumar V, Ross Quinlan J, et al. Top 10 algorithms in data mining. *Knowl Inf Syst*. 2008;14(1):1-37. doi:10.1007/s10115-007-0114-2
101. Decision tree learning - Wikipedia. Accessed April 12, 2023. https://en.wikipedia.org/wiki/Decision_tree_learning#/media/File:Decision_Tree.jpg
102. Tin Kam Ho. Random decision forests. In: *Proceedings of 3rd International Conference on Document Analysis and Recognition*. IEEE Comput. Soc. Press; 1995:278-282. doi:10.1109/ICDAR.1995.598994
103. Breiman L. Random Forests. *Springer Science and Business Media LLC*. 2001. doi:10.1023/a:1010933404324
104. Random forest - Wikipedia. Accessed June 12, 2023. https://en.wikipedia.org/wiki/Random_forest#/media/File:Random_forest_diagram_complete.png
105. Liaw A, Wiener M. Classification and Regression by randomForest. *R News*. 2002;2(3):18-22.
106. Pirayonesi SM, El-Diraby TE. Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index. *J Infrastruct Syst*. 2020;26(1). doi:10.1061/(ASCE)IS.1943-555X.0000512
107. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Statist*. 2001;29(5):1189-1232. doi:10.1214/aos/1013203451

108. Powers D. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *Mach Learn Technol.* 2008;2.
109. Altman DG, Bland JM. Statistics Notes: Diagnostic tests 1: sensitivity and specificity. *BMJ.* 1994;308(6943):1552-1552. doi:10.1136/bmj.308.6943.1552
110. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging.* 2015;15(1):29. doi:10.1186/s12880-015-0068-x
111. Mohammed R, Rawashdeh J, Abdullah M. Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. In: *2020 11th International Conference on Information and Communication Systems (ICICS).* IEEE; 2020:243-248. doi:10.1109/ICICS49469.2020.239556
112. Ling CX, Li C. Data Mining for Direct Marketing: Problems and Solutions. In: *Knowledge Discovery and Data Mining.* ; 1998.
113. Fernandez A, Garcia S, Herrera F, Chawla NV. SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. *jair.* 2018;61:863-905. doi:10.1613/jair.1.11192
114. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning.* 2nd ed. Springer New York; 2009:106-119. doi:10.1007/978-0-387-84858-7
115. Little R, Rubin D. *Statistical Analysis with Missing Data, Third Edition.* Wiley; 2019. doi:10.1002/9781119482260
116. Portin P. The development of genetics in the light of thomas kuhn's theory of scientific revolutions. *Recent Adv DNA Gene Seq.* 2015;9(1):14-25. doi:10.2174/2352092209666150921110920
117. Mendel G. *Versuche Über Pflanzen-Hybriden /.* Im Verlage des Vereines,; 1866. doi:10.5962/bhl.title.61004
118. Johannsen W. Elemente der exakten Erblchkeitslehre. Mit Grundzügen der biologischen Variationsstatistik. *ZVer-erbungslehre.* 1914;11(1):200-200. doi:10.1007/BF01704312
119. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* 1953;171(4356):737-738. doi:10.1038/171737a0
120. Human Genome Project Fact Sheet. Human Genome Project Fact Sheet. Accessed May 1, 2023. <https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project>

121. Saiki RK, Scharf S, Faloona F, et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. 1985;230(4732):1350-1354. doi:10.1126/science.2999980
122. Saiki RK, Gelfand DH, Stoffel S, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. 1988;239(4839):487-491. doi:10.1126/science.239.4839.487
123. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*. 1977;74(12):5463-5467. doi:10.1073/pnas.74.12.5463
124. International Human Genome Sequencing Consortium, Whitehead Institute for Biomedical Research, Center for Genome Research., Lander ES, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062
125. Smith DR, Quinlan AR, Peckham HE, et al. Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res*. 2008;18(10):1638-1642. doi:10.1101/gr.077776.108
126. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*. 2010;11(1):31-46. doi:10.1038/nrg2626
127. Mardis ER. A decade's perspective on DNA sequencing technology. *Nature*. 2011;470(7333):198-203. doi:10.1038/nature09796
128. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*. 2009;461(7261):272-276. doi:10.1038/nature08250
129. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*. 2011;12(11):745-755. doi:10.1038/nrg3031
130. Jiang Y, Yuen RKC, Jin X, et al. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. *Am J Hum Genet*. 2013;93(2):249-263. doi:10.1016/j.ajhg.2013.06.012
131. Shah PD, Nathanson KL. Application of Panel-Based Tests for Inherited Risk of Cancer. *Annu Rev Genomics Hum Genet*. 2017;18:201-227. doi:10.1146/annurev-genom-091416-035305
132. Genetic Testing Techniques - Testing.com. Genetic Testing Techniques - Testing.com. Accessed May 2, 2023. <https://www.testing.com/genetic-testing-techniques/>
133. Langer-Safer PR, Levine M, Ward DC. Immunological method for mapping genes on Drosophila polytene chromosomes. *Proc Natl Acad Sci USA*. 1982;79(14):4381-4385. doi:10.1073/pnas.79.14.4381

134. Taub FE, DeLeo JM, Thompson EB. Sequential comparative hybridizations analyzed by computerized image processing can identify and quantitate regulated RNAs. *DNA*. 1983;2(4):309-327. doi:10.1089/dna.1983.2.309
135. Chapman S, Saladi S, Baw M. rcpch/rcpchgrowth-python: v3.1.17. *Zenodo*. 2022. doi:10.5281/zenodo.6301089
136. Bayram RO, Özdemir H, Emsen A, Türk Dağı H, Artaç H. Reference ranges for serum immunoglobulin (IgG, IgA, and IgM) and IgG subclass levels in healthy children. *Turk J Med Sci*. 2019;49(2):497-505. doi:10.3906/sag-1807-282
137. *Lanzkowsky's Manual of Pediatric Hematology and Oncology*. Elsevier; 2022. doi:10.1016/C2019-0-03654-0
138. Besci Ö, Başer D, Ögülür İ, et al. Reference values for T and B lymphocyte subpopulations in Turkish children and adults. *Turk J Med Sci*. 2021;51(4):1814-1824. doi:10.3906/sag-2010-176
139. Python 3.10. Accessed May 24, 2023. <https://www.python.org/>
140. The pandas development team. pandas-dev/pandas: Pandas 1.0.3. *Zenodo*. 2020. doi:10.5281/zenodo.3509134
141. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12:2825-2830.
142. Lemaitre G, Nogueira F, Aridas CK. Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *Journal of Machine Learning Research*. 2017;18(17):1-5.
143. Plotly: Low-Code Data App Development. Accessed April 12, 2023. <https://plotly.com/>
144. Kammermeier J, Lamb CA, Jones KDJ, et al. Genomic diagnosis and care co-ordination for monogenic inflammatory bowel disease in children and adults: consensus guideline on behalf of the British Society of Gastroenterology and British Society of Paediatric Gastroenterology, Hepatology and Nutrition. *Lancet Gastroenterol Hepatol*. 2023;8(3):271-286. doi:10.1016/S2468-1253(22)00337-5
145. Akkelle BS, Tutar E, Volkan B, et al. Gastrointestinal Manifestations in Children with Primary Immunodeficiencies: Single Center: 12 Years Experience. *Dig Dis*. 2019;37(1):45-52. doi:10.1159/000492569