DOI: 10.4274/forbes.galenos.2023.61687
Forbes J Med 2023;4(1):95-102

Orijinal Araştırma / Original Article

Makine Öğrenmesi Yönteminin Laboratuvar Tıbbına Uygulama Örneği: Yenidoğan Referans Aralıklarının Belirlenmesi

An Example of Application of Machine Learning Method to Laboratory Medicine: Determination of Newborn Reference Ranges

Oktay YILDIRIM[1], Özlem AKTAŞ[2], Ali Rıza ŞİŞMAN[5], Dilek ORBATU[4], Senem ALKAN ÖZDEMİR[4], Adem AYDIN[6], Eminullah YAŞAR[1], Mohammed Abebe YİMER[7], Süleyman SEVİNÇ[3]

[1]Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, İzmir, Türkiye
[2]Dokuz Eylül Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Bilgisayar Donanımı Anabilim Dalı, İzmir, Türkiye
[3]Dokuz Eylül Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Bilgisayar Yazılımı Anabilim Dalı, İzmir, Türkiye
[4]Sağlık Bilimleri Üniversitesi, İzmir Dr. Behçet Uz Çocuk Hastalıkları ve Cerrahisi Eğitim ve Araştırma Hastanesi, İzmir, Türkiye
[5]Dokuz Eylül Üniversitesi Tıp Fakültesi, Tıbbi Biyokimya Anabilim Dalı, İzmir, Türkiye
[6]Dokuz Eylül Üniversitesi Tıp Fakültesi, Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, İzmir, Türkiye
[7]Arba Minch Üniversitesi, Bilgisayar Mühendisliği Bölümü, Arba Minch, Etiyopya

Atıf: Yıldırım O, Aktaş Ö, Şişman AR, Orbatu D, Alkan Özdemir S, Aydın A, Yaşar E, Yimer MA, Sevinç S. An Example of Application of Machine Learning Method to Laboratory Medicine: Determination of Newborn Reference Ranges. Forbes J Med 2023;4(1):95-102

ÖZ

Amaç: Konvansiyonel yöntemlerle referans aralıklarının belirlenmesindeki zorluklardan dolayı, çağdaş makine öğrenmesi yöntemlerini kullanarak referans aralıklarının belirlenmesi amaçlanmıştır.

Yöntem: 2018-2019-2020 yıllarında Dokuz Eylül Üniversitesi Merkez Laboratuvarı'nda çalışılan, yenidoğan dönemine ait inorganik fosfor, kalsiyum, kreatinin, neonatal bilirubin, üre azotu testlerine ait sonuçlar hastane veri tabanından alındı. Geliştirdiğimiz denetimsiz makine öğrenmesi algoritması yardımıyla ilgili testlere ait yaş kırılımları ve bu bağlı referans aralıkları hesaplandı.

Bulgular: Geliştirdiğimiz denetimsiz makine öğrenme yönteminin referans aralıklarının belirlenmesinde indirekt yöntemlere yeni, çağdaş bir alternatif olduğu belirlendi. Bu yöntemle test sonuçlarının anlamlı değişkenlik gösterdiği yaş aralıkları yüksek çözünürlükte teste özgü şekilde bulundu.

Sonuç: Bilgisayar işlemci gücündeki artışlar, yeni özgün yapay zeka ve makine öğrenmesi temelli algoritmalar, yüksek miktarda veri depolamayı sağlayan veri tabanları, referans aralıklarının belirlenmesi için çağdaş bir çözüm geliştirme imkanı sunmaktadır. Bu çalışmada, referans aralıklarının hesaplanmasında temel basamak olan yaş aralıklarını yüksek çözünürlükte belirleyebilen, matematiksel ve istatistiksel temellere dayanan, denetimsiz makine öğrenme algoritması çözümü sunulmuştur. Çalışmada geliştirilen algoritmik yöntemi kullanarak her laboratuvar kendi popülasyon ve analitik yöntemlerine uygun referans aralıklarını kolay, hızlı, güvenli ve ekonomik bir şekilde hesaplayabilecektir.

Anahtar Kelimeler: Makine öğrenmesi, referans aralıkları, yenidoğan, elektronik hasta kayıtları, normal dağılımların örtüşmesi

ABSTRACT

Objective: Due to the difficulties in determining reference intervals with conventional methods, it determines them using modern machine learning methods.

Methods: The results of the newborns' inorganic phosphorus, calcium, creatinine, neonatal bilirubin, and urea nitrogen tests, which were studied in the Dokuz Eylül University Central Laboratory for the years 2018-2019-2020, were obtained from the hospital database. The unsupervised machine learning algorithm we developed calculated test-specific age intervals and related reference intervals.

Geliş/Received: 01.12.2022
Kabul/Accepted: 04.01.2023

Sorumlu Yazar/ Corresponding Author:
Dr. Oktay YILDIRIM
Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, İzmir, Türkiye
Tel.: +90 544 777 48 66
oktay.yildirim@deu.edu.tr
ORCID: 0000-0001-7281-3623

©Telif Hakkı 2023 İzmir Buca Seyfi Demirsoy Eğitim ve Araştırma Hastanesi / Forbes Tıp Dergisi, Galenos Yayınevi tarafından yayınlanmıştır.
Bu dergide yayınlanan bütün makaleler Creative Commons 4.0 Uluslararası Lisansı (CC-BY) ile lisanslanmıştır.

©Copyright 2023 by the İzmir Buca Seyfi Demirsoy Training and Research Hospital / Forbes Journal of Medicine published by Galenos Publishing House.
Licensed by Creative Commons Attribution 4.0 International (CC BY)

Results: It was determined that the unsupervised machine learning method we developed is a new, contemporary alternative to indirect methods for determining reference intervals. With this method, the age ranges in which the test results showed significant variability were found in a high-resolution test-specific manner.

Conclusion: Increases in computer processing power, new original artificial intelligence and machine learning-based algorithms, and databases that store large amounts of data offer a contemporary solution for determining reference intervals. In this study, an unsupervised machine learning algorithm solution based on mathematical and statistical foundations, which can determine age ranges, which is the basic step in the calculation of reference intervals, with high resolution is presented. By using the algorithmic method developed in the study, each laboratory will be able to calculate reference intervals suitable for their population and analytical methods in an easy, fast, safe, and economical way.

Keywords: Machine learning, reference intervals, newborn, electronic patient records, overlapping normal distributions

GİRİŞ

Sağlıklı bireylerin biyokimyasal değerlerinin dağılımı referans aralıkları olarak ifade edilir. Referans aralıkları kavramı ilk olarak 1969 yılında Klinik Laboratuvar Tıbbı Kongresi'nde Grasbeck ve Saris[1] tarafından ortaya atılmıştır. Referans aralıklarının belirlenmesi klinik biyokimya alanında çok önemlidir. Laboratuvar test sonuçları, hekimlerin tanı ve tedavi süreçlerine yardımcı olacak nesnel veriler sağlar. Referans aralıklarının belirlenmesine ilişkin tavsiyeler, Uluslararası Klinik Kimya Federasyonu (IFCC) tarafından altı bölüm halinde yayınlanmıştır.[2-7] Bu altı parçalı referans aralığı belirleme tavsiyelerinde IFCC, bölgesel farklılıklar, popülasyon farklılıkları ve laboratuvar farklılıkları nedeniyle her laboratuvarın kendi referans aralığını belirlemesini önerir. Ancak Clinical and Laboratory Standards Institute (CLSI) yönergelerinde belirttiği klasik yöntemlerle her laboratuvarın kendi referans aralığını belirlemesi çok maliyetli ve zordur.[8] CLSI referans aralığı belirlemek için her bir alt gruba ait en az 120 referans birey olmasını önermektedir. Bu bireylerin akut ya da kronik bir rahatsızlığının olamaması, herhangi bir ilaç, alkol, sigara kullanmıyor olması gerekmektedir.[3] Özellikle pediatrik grupta ebeveynlerden onam alınmasının güçlüğünden dolayı gönüllü referans bireyin bulunması oldukça zordur.

Hali hazırda kullanılan referans aralıklarının büyük çoğunluğu erişkin yaşları kapsamaktadır. Çocukluk çağına ait referans aralıkları çok az ve kısıtlıdır. Birçok kit üreticisi kit prospektüsünde çocukluk dönemi için referans aralığı bilgisi vermemektedir ya da kısıtlı bir yaş aralığı için vermektedir. Çocuklar için daha çok klinik ya da laboratuvar textbooklarda yer alan değerler kullanılmaktadır. Çocukluk dönemi referans aralıklarının belirlenmesi konusunda The Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER)[9] başta olmak üzere değişik çalışma grupları çalışmalar yürütmektedir. Ancak bu çalışmalar uzun yıllar alan maliyetli çalışmalardır. Ayrıca buradan çıkan verilerin küresel olarak başka toplumlara, farklı kit ve analizörlerden elde edilen sonuçlara uygulanması uygun olmayabilir.

Maalesef çocukluk yaş dönemine ait referans aralıklarının bulunmamasından dolayı laboratuvarlar test sonuç raporlarında çocuklar için referans aralığı alanını boş bırakmakta ya da erişkin referans aralıkları ile rapor etmektedir. Bu durum klinisyenler için yanıltıcı veya eksik bilgi oluşturmakta, teşhis ve tedavi süreçlerini olumsuz yönde etkileyebilmektedir.

Bir teste ait referans aralıkları genellikle yaşa ve cinsiyete göre gruplandırılarak yayınlanır. Hali hazırda yapılan çocukluk çağına ait referans aralığı çalışmalarında yaş aralığı kırılımları gelişme evreleri dikkate alınarak ayrıştırılmakta ya da tüm yaşlar birlikte tek bir yaş grubu olarak ele alınmaktadır. Çocukluk döneminin tamamını tek bir referans aralığıyla değerlendirmek birçok test için doğru değildir. Hatta gelişme evrelerine göre yapılan yaş kırılımları da yetersiz olabilmektedir. CLSI, her testin fizyolojik özelliklerine göre farklı alt yaş kırılımlarının olabileceğini, bu nedenle alt kırılımların teste spesifik belirlenmesi gerektiğini belirtmektedir.[8] Alt kırılımların birbirlerinden farklı olup olmadığını belirlemek için Harris ve Boyd[10] v ve Lahti ve ark.'nın[11,12] önerdiği yöntemler bulunmaktadır. Ancak bu konuda bir konsensüs bulunmamakta olup, CLSI bu alanda daha ileri çalışmalara gereksinim duyulduğunu belirtmektedir.[8]

Makine öğrenimi, bilgisayar programlarının deneyim yoluyla otomatik olarak gelişmesini sağlayan ve belirli verilerden genel yasaları otomatik olarak çıkaran algoritmaları inceleyen bilgisayar bilimi alanıdır.[13] Makine öğrenimi ve yapay zeka kavramları sıklıkla birlikte kullanılır ancak makine öğrenimi yapay zekanın bir alt dalıdır. Makine öğrenimi algoritmaları, istatistiksel ve matematiksel yöntemler kullanarak sınıflandırmalar veya tahminler yapar. Makine öğrenimi algoritmaları, denetimli (supervised), denetimsiz (unsupervised) ve pekiştirmeli (reinforcement) olarak üçe ayrılır. Denetimli öğrenme algoritmaları hem girdileri hem de istenen çıktıları içeren bir veri kümesinin matematiksel bir modelini oluşturur[14]. Bu makine öğrenimi paradigmasında, modeli eğitmek için kullanılan eğitim verilerinin önceden etiketlenmesi gerekir. Denetimsiz makine öğrenimi paradigması, yalnızca girdilerden oluşan etiketlenmemiş verileri kullanır. Denetimsiz makine öğrenimi algoritmaları, insan müdahalesi olmadan gizli kalıpları veya veri gruplarını keşfeder. Veri noktalarının uzaklıklarını, komşuluk ilişkilerini ve yoğunluklarını kullanarak veriler hakkında çıkarımlarda bulunur.

"Principal Component Analysis", "Hierarchical", "K-means" ve "Gaussian Mixture Model" (GMM) sıklıkla kullanılan denetimsiz makine öğrenmesi algoritmalarıdır. Pekiştirmeli öğrenme ise, deneyimlerden öğrenerek ve performansını iyileştirerek çevresini otomatik olarak keşfeden, geri bildirime dayalı bir öğrenme paradigmasıdır. Pekiştirmeli öğrenmede, denetimli öğrenmedeki gibi etiketli veri yoktur ve sadece deneyimden öğrenilir.

Bu çalışmada temel amaç, makine öğrenmesi algoritmalarının laboratuvar tıbbına uygulanarak referans aralıklarının hesaplanabileceğini göstermektir. Bu amaçla denetimsiz makine öğrenmesi algoritması ile hastane verileri kullanılarak her hastanenin kendi referans aralığını belirleyebileceği çağdaş bir yöntem sunulmuştur. Çalışmanın ilk aşamasında yaş aralığı kırılımlarının denetimsiz makine öğrenmesiyle otomatik olarak belirlenmesi, daha sonra her bir yaş aralığı kırılımı için referans aralıklarının hesaplanması hedeflenmiştir.

YÖNTEM

2018-2019-2020 yıllarında Merkez Laboratuvarı'nda çalışılan, 1-28 günlük bebeklere ait laboratuvar test sonuçları hastane veri tabanından alınarak retrospektif olarak incelendi. Yenidoğan döneminde bebeklerde sıklıkla takip edilen beş test seçildi. Bunlar; serumda inorganik fosfor, kalsiyum, kreatinin, üre azotu ve plazmada neonatal bilirubindir. Laboratuvar bilgi yönetim sistemi veri tabanından laboratuvar uzmanı tarafından onaylanmış kliniğe rapor edilmiş veriler, hasta kimlik bilgileri anonimize edilerek; hastanın doğum tarihi, testin yapılış tarihi, hasta cinsiyeti, test adı, test sonucu ve test birimi olarak alındı. Rapor edilebilir aralık dışındaki sonuçlar, yaşamla bağdaşmayan uç sonuçlar çalışmadan çıkarıldı. Klinik pratikte yenidoğan dönemde cinsiyet farklılığı dikkate alınmadığından erkek ve kızlara ait sonuçlar birlikte değerlendirildi. Yenidoğan döneminde annenin kadın doğum servisindeki yatış kaydından dolayı, bebekler de bilgi sisteminde yatan hasta olarak görünmektedir. Bu sebeple yenidoğan dönemi hem servis hem de poliklinik başvuruları içerdiğinden, sonuçlarda servis poliklinik ayrımı yapılmadı.

İnorganik fosfor, kalsiyum, kreatinin, üre azotu testlerine ait sonuçlar orijinal kitiyle Beckman-Coulter AU5800 (Japon) otoanalizöründe üretilmiştir. Neonatal bilirubin testine ait sonuçlar Apel BR 5000 N (Japon) bilirubinmetre cihazında üretilmiştir. Çalışmayı kapsayan üç yıllık süreçte cihaz ve yöntemlerde değişiklik olmamıştır.

Klinik pratikte referans aralıkları yaş ve cinsiyet farklılıkları dikkate alınarak kullanılmaktadır. Yaş ve cinsiyete göre alt popülasyonları doğru belirlemek referans aralığı çalışmalarının kritik adımıdır. Çalışmamızda alt popülasyonların belirlenmesi bir kümeleme problemi olarak ele alınmıştır. Kümeleme problemleri denetimsiz makine öğrenimi algoritmalarıyla etkili bir şekilde çözülebilmektedir. "K-means", aynı kümedeki noktalar arasındaki ortalama mesafenin karesini en aza indirmeyi amaçlayan, yaygın olarak kullanılan bir kümeleme algoritmasıdır.[15] GMM, K-means algoritmasının daha sofistike bir versiyonudur. GMM'de verilerin "k" adet normal dağılımın karşımı olduğu varsayılır.[16] Her bir kümenin normal dağılım özelliği göstermesi referans aralığı verileri için de uygundur. Ancak "K-means" ve GMM algoritmalarında başlangıçta kaç tane küme aranacağı seçilmek zorundadır. "K-means" ve GMM ile ilgili başka bir kısıt ise aynı yaşa ait sonuçların farklı kümelere ait olabilmesidir. Bu durum, referans aralıklarının klinik pratikteki kullanımına uygun değildir. Bir diğer kümeleme algoritması olan Hiyerarşik Kümeleme, başlangıçta her veri noktasını ayrı bir küme olarak ele alıp, her adımda benzer nesneleri gruplayarak ilerleyen bir algoritmadır. Gruplandırma, her grubun diğer gruptan farklı olduğu ve her gruptaki nesnelerin birbirine benzediği aşamada sona erer.[17] Başlangıçta küme sayısı belirtme zorunluğunun olmaması ve kümelerin kırılımlarının yaşa göre belirlenmesine imkan sağlamasından dolayı çalışmamızda Hiyerarşik Kümeleme Algoritması tercih edilmiştir. Algoritmanın adımları aşağıdaki gibidir:

Adım 1: Sonuçları yaş bazında maksimum sayıda kümeye ayır.

Adım 2: Yüz yirmiden az sonuç olan kümeyi bir sonraki kümeyle birleştir.

Adım 3: Her küme için Tukey[18] metodunu kullanarak uç değerleri temizle.

Adım 4: Her küme için Kolmogorov-Smirnov yönetemi[19] kullanılarak normalite testi yap.

Adım 5: Tüm kümeler normal dağılım gösteriyorsa Adım 7'ya geç. Normal dağılmayan kümelere Box-Cox[20] transformasyonu uygula, Tukey metoduyla uç değerleri temizle Kolmogorov-Smirnov yönetemi kullanılarak normalite testi yap.

Adım 6: Tüm kümeler normal dağılım gösteriyorsa Adım 7'ya geç. Normal dağılmayan kümeler varsa veriler uygun değil. Algoritmayı sonlandır.

Adım 7: Her kümenin ortalama ve standart sapma değerlerini hesapla.

Adım 8: İlk günden başlayarak son güne kadar her bir kümenin bir sonraki kümeyle benzerliğini Normal Dağılımların Örtüşmesi yöntemiyle[21] hesapla.

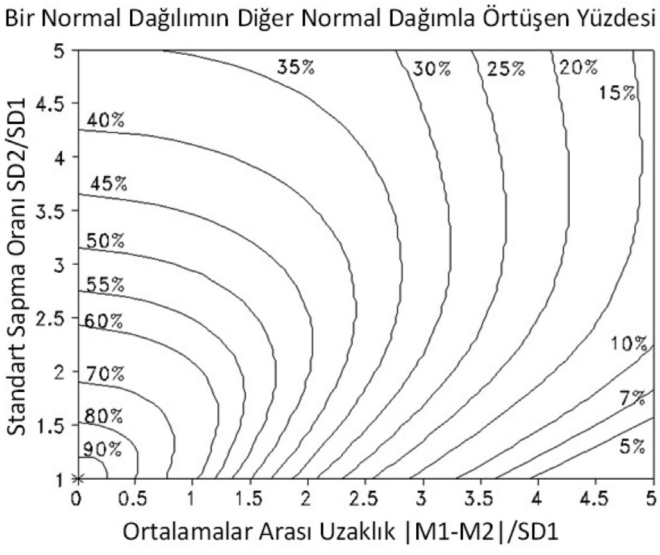
Adım 9: Maksimum benzerlik gösteren ardışık iki kümeyi belirle.

Adım 10: Maksimum benzerlik gösteren ardışık iki kümenin benzerliği %95'ten yüksekse bu iki kümenin sonuçlarını birleştir. Yeni kümenin ortalama ve standart sapma değerlerini hesapla.

Adım 11: Adım 8, 9, 10'u benzerlik %95'ten yüksek olduğu sürece tekrar et. Aksi durumda algoritmayı sonlandır. Bulunan kümeleri referans aralığının yaş kırılımı olarak kabul et.

Adım 12: Her küme için alt sınır 2,5; persentil, üst sınır 97,5. persentil olacak şekilde referans aralıklarını belirle.

Algoritmanın birinci adımında, çalışmamızda 0-28 günlük bebeklere ait sonuçları kullandığımız için en yüksek yaş çözünürlüğü 1 gün olarak kabul edilmiştir. Bu yüzden kümeler gün gün olacak şekilde ayrılmıştır. Üçüncü adımdaki Tukey metodu, dördüncü adımdaki Kolmogorov-Smirnov metodu ve beşinci adımdaki Box-Cox metodu için, python scipy.stats[22] kütüphanesi kullanılmıştır. Sekizinci adımda kullanılan Normal Dağılımların Örtüşmesi yöntemi, iki normal dağılım eğrisinin üst üste örtüşme yüzdesini hesaplar. Bu yöntemi uygulayabilmek için iki dağılımın ortalamaları (M1 ve M2) ve standart sapmaları (SD1 ve SD2) bulunur. Günlerden daha küçük standart sapmaya sahip olan SD1, diğeri ise SD2 olarak etiketlenir. Günler arasındaki örtüşmenin yüzdesini bulmak için Şekil 1'de yer alan nomogram kullanılır. Bu nomogramda x-ekseninde ortalamalar arasındaki normalize edilmiş mutlak mesafe (|M2- M1|/SD1), y-ekseninde ise standart sapmaların oranı (SD2/SD1) yer almaktadır. Nomogram kullanılarak iki kırılım arasındaki üst üste örtüşmenin oranı yüzde olarak ifade edilir.[21] Normal Dağılımların Örtüşmesi



Şekil 1. Normal dağılımların örtüşmesi nomogram

hesabı için "python statistics" kütüphanesinin "overlap" fonksiyonu kullanılmıştır.[22,23] Başlangıçta her kümenin normal dağıldığını kontrol ettiğimiz için, onuncu adımdaki birleştirme işlemi sonrasında ortaya çıkan yeni dağılımın, Cramer'in normal dağılım için ayrıştırma teoremine göre normal dağılması beklenir. Cramer'in teoremi, bağımsız normal dağılımlı rasgele değişkenler ε1, ε2 verildiğinde, bunların toplamlarının da normal dağılım gösterdiğini söyler.[24]

Bu algoritma Python 3.9 programlama dili kullanılarak geliştirilmiştir. Intel(R) Core(TM) i7-10750H 2.60GHz işlemci, 16 GB RAM, 256 GB SDD sabit disk donanımı üzerinde her bir testin referans aralığı ortalama 3,5 dakikada hesaplanmıştır.

BULGULAR

İnorganik fosfor testi için 1892 erkek, 1440 kız yenidoğana ait toplam 3332 test sonucu algoritmaya sunuldu. Algoritma, test sonuçlarının 1-28. gün olarak tek bir yaş kırılımına ait olduğunu belirledi. Bu yaş aralığına ait referans aralıkları Tablo 1'de verilmiştir.

Kalsiyum testi için 2428 erkek, 1849 kız yenidoğana ait toplam 4277 test sonucu algoritmaya sunuldu. Algoritma, test sonuçlarının 1-28. gün olarak tek bir yaş kırılımına ait olduğunu belirledi. Bu yaş aralığına ait referans aralıkları Tablo 1'de verilmiştir.

Kreatinin testi için 2270 erkek, 1741 kız yenidoğana ait toplam 4011 test sonucu algoritmaya sunuldu. Algoritma, test sonuçlarının 1-2. gün, 3-5. gün, 6-11. gün, 12-14. gün, 15-16. gün, 17-24. gün, 25-28. gün olarak 7 farklı yaş kırılımına ait olduğunu belirledi. Bu yaş aralıklarıma ait referans aralıkları Tablo 1'de verilmiştir.

Neonatal bilirubin testi için 3782 erkek, 2808 kız yenidoğana ait toplam 6590 test sonucu algoritmaya sunuldu. Algoritma, test sonuçlarının 1. gün, 2. gün, 3. gün, 4-28. gün olarak 4 farklı yaş kırılımına ait olduğunu belirledi. Bu yaş aralıklarına ait referans aralıkları Tablo 1'de verilmiştir.

Üre azotu testi için 2227 erkek, 1719 kız yenidoğana ait toplam 3946 test sonucu algoritmaya sunuldu. Algoritma, test sonuçlarının 1. gün, 2-11. gün, 12-26. gün, 27-28. gün olarak 4 farklı yaş kırılımına ait olduğunu belirledi. Bu yaş aralıklarına ait referans aralıkları Tablo 1'de verilmiştir.

TARTIŞMA

Laboratuvar sonuçları hastalıkların değerlendirilmesinde tıpta çok önemli bir yere sahiptir. Ancak test sonuçları tek başına yeterli bilgiyi sağlamamaktadır. Bu nedenle sonuçlar referans aralıkları ile birlikte rapor edilir. Referans aralıkları, hastanın yaş, cinsiyet gibi demografik bilgileri ile eşleştirilerek rapor edilmelidir.[25] Farklı demografik özelliklere sahip popülasyonlar için ayrı ayrı referans aralığı belirleme çalışmaları yapılmalıdır.

Bunun için öncelikle alt popülasyonlar belirlenmeli daha sonra her popülasyona ait en az 120 sağlıklı bireyden alınan örneklerin çalışılması gerekmektedir. Direkt yöntem olarak bilinen bu çalışmayı her test için yapmak oldukça zor ve maliyetlidir. Direkt yöntem kullanılarak yapılan referans aralığı çalışmaları daha çok erişkin popülasyona ait olup, pediatrik yaş grupları için referans aralığı çalışmaları oldukça kısıtlı sayıdadır. Bu durum özellikle pediatrik hastaların sonuçlarının değerlendirilmesinde eksiklik ve hatalara sebep olabilecektir. Pediatrik yaş grupları içerisinde de özellikle yenidoğan dönemi için yapılan referans aralığı çalışmaları daha da azdır. Klinik pratikte sıklıkla Nelson Textbook of Pediatrics[26] (Nelson) kitabının yayınladığı referans aralıklarının kullanıldığı görülmektedir. Yayınlanmış pediatrik referans aralıklarının pratikte kullanımındaki temel sorun, sonuçların farklı toplum, etnisite, analizör ve kitlerden elde edilmesinden dolayı küresel olarak kullanılamamasıdır. Tablo 2'de Nelson kitabında, Tablo 3'te CALIPER çalışmasında verilen yaş aralıkları ve referans değerlerin itibariyle birbirinden farklılaştığı görülmektedir. Bu sorunlardan dolayı IFCC her laboratuvarın kendi referans aralıklarını belirlemesini tavsiye etmektedir. Geliştirdiğimiz çözüm her laboratuvarın geçmiş verilerini kullanarak, kendi hasta popülasyonuna, kendi analizör ve kitine uygun referans aralığını kolaylıkla belirlemesine imkan sağlayabilecektir.

Referans aralığı belirlemek için altın standart olarak kullanılan direkt yöntemin zorluklarından dolayı, hastane veya laboratuvar veri tabanından elde edilen retrospektif veriler ile referans aralıklarının belirlenmesi için indirekt yöntem kullanılmaktadır. İndirekt yönetim kullanımında dikkat edilecek hususlar IFCC tarafından rapor edilmiştir.[27] Yapmış olduğumuz çalışma da indirekt referans aralığı belirleme yöntemlerinden biridir. İndirekt yöntemlerin direkt yöntemlere göre temel avantajları, süreçlerin daha hızlı ve ucuz olması, bireylere invaziv bir girişim gerektirmemesidir, ancak veriler içerisinde hasta kişilere ait sonuçların olması bir kısıttır. Tanı kodları, patolojik sonuçları belirlemek için kullanılan araçlardan biridir.

Tablo 1. Denetimsiz makine öğrenmesi ile belirlenen yenidoğan yaş kırılımları ve referans aralıkları

	Günler									
Testler	**1**	**2**	**3**	**4-5**	**6-11**	**12-14**	**15-16**	**17-24**	**25-26**	**27-28**
Fosfor, inorganik [serum] (mg/dL)	3,04-8,41									
Kalsiyum [serum] (mg/dL)	7,03-11,31									
Kreatinin [serum] (mg/dL)	0,44-1,32		0,33-1,38		0,23-1,38	0,22-1,25	0,19-1,39	0,16-1,11	0,14-0,92	
Neonatal bilirubin [plazma] (mg/dL)	2,68-13,31	3,7-14,9	4-18,72	4,03-18,7						
Üre azotu [serum] (mg/dL)	4,79-28,15	3,17-34,6				2,9-34,7				2,64-23,8

Tablo 2. Nelson yenidoğan referans aralıkları

	Günler					
Testler	**1**	**2**	**3**	**4-5**	**6-7**	**8-28**
Fosfor, inorganik [serum] (mg/dL)	4,8-8,2				N/A*	
Kalsiyum [serum] (mg/dL)	9-10,6	7-12	N/A*	9-10,9		N/A*
Kreatinin [serum] (mg/dL)	0,3-1					
Neonatal bilirubin [plazma] (mg/dL)	<6	<8	<12		<10	
Üre azotu [serum] (mg/dL)	3-12					

*N/A: Yayınlanmamış

Tablo 3. CALIPER yenidoğan referans aralıkları		
	Günler	
Testler	**1-14**	**15-28**
Fosfor, inorganik [serum] (mg/dL)	5,57-10,52	4,77-8,42
Kalsiyum [serum] (mg/dL)	8,54-10,98	
Kreatinin [serum] (mg/dL)	0,42-1,05	0,32-0,53
Neonatal bilirubin [plazma] (mg/dL)	0,19-16,6	0,05-0,68
Üre azotu [serum] (mg/dL)	2,8-22,96	3,36-16,8
CALIPER: The Canadian Laboratory Initiative on Pediatric Reference Intervals		

Ancak ülkemizde bu kodlar hastalığı tanımlamaktan ziyade faturalama ve test istemini yapabilmek için ön koşul olarak kullanılmaktadır. Bu durum hatalı hastalık kodlarının seçilmesine sebep olmaktadır. Çalışmamızda da doğru hastalık kodları seçilmediği için patolojik sonuçlar dışlanamamıştır. Patolojik sonuçlar algoritmanın çalışma prensibinde değişikliğe neden olmamıştır, ancak geliştirilen yöntemin klinik çalışmaları yapılırken indirekt yöntemler için önerilen dışlama kriterlerinin uygulanması daha uygun olacaktır.

Gerek direkt gerekse indirekt yöntemlerin ortak eksikliği referans aralığı belirlenecek alt popülasyonlarının yaş aralıklarının seçiminin zorluğudur. CLSI her testin kendine özgü fizyolojik özelliğinden dolayı farklı yaş kırılımlarına sahip olduğunu, yaş kırılımlarının teste özel belirlenmesi gerektiğini belirtmektedir.[8] Yani tüm testlere genellenebilecek ortak bir yaş kırılımı varsayımı yapmak doğru değildir. Örneğin; Tablo 3'te yer alan CALIPER çalışmasına ait referans aralıkları incelendiğinde, neotanal dönemde tüm testlerin 1-14. gün ve 15-28. gün olarak sabit bir yaş aralığı ile yayınlandığı görülmektedir. Bu durum gerçek yaşamla bağdaşmamaktadır. Oysa ki Nelson'ın yayınladığı referans aralıklarında her teste özgü farklı yaş kırılımlarının olduğu Tablo 2'de görülmektedir. Çalışmalar arasındaki bu farklılıklar referans aralıklarının etkin kullanılmasında sorun yaratmaktadır. Çalışmamızda, bu eksikliği gidermek için hastane verileri üzerinde denetimsiz makine öğrenmesi yöntemiyle yaş kırılımlarının doğru çözünürlükte belirleyen bir algoritma geliştirilmiştir. Böylece her teste özgü, testin sonuçlarında anlamlı değişikliklerin olduğu gerçek yaşamla uyumlu yaş aralıkları belirlenebilmiştir. Örneğin Tablo 1'de kreatinin, üre azotu ve neonatal bilirubin testleri birden fazla yaş kırılımına sahipken, inorganik fosfor ve kalsiyum testleri ise tek bir yaş kırılımına sahiptir. Yaş kırılımındaki bu farklı çözünürlüklerin etkinliğinin doğrulanması için retrospektif ya da prospektif klinik çalışmalara ihtiyaç vardır.

Klinik laboratuvarlarda daha çok kit ve analitik sistem üreticilerinin tedarik ettiği referans aralıkları yaygın olarak kullanılmaktadır. Üreticiler bunu kendi yaptıkları çalışmalara ya da literatürdeki çalışmalara dayandırmaktadırlar. Her iki durumda da bazı yaş aralıkları için referanslar verilmemektedir, yani boşluklar bulunmaktadır. Bu durumda laboratuvarlar üç farklı çözümle bu eksikliği gidermeye çalışmaktadırlar. Bunlar; ilgili yaş aralığına ait referans değerin rapor edilmemesi, farklı bir yaş aralığına ait değerin rapor edilmesi ve farklı kaynaklardaki referans aralıkları ile kombine kullanılması şeklinde karşımıza çıkmaktadır. Ancak bu çözümlerin tamamı sonuçların doğru ve etkin değerlendirilmesinde uygunsuzluk oluşturmaktadır. Referans aralıklarındaki boşluklar ve uygunsuzluklar, hekimlerin kafa karışıklığına sebep olmasının yanında, online ya da çıktı olarak sonuç raporuna bakan ebeveynlerde sonuçların normal olup olmadığı hakkında merak, endişe ve paniğe sebep olabilmektedir. Örneğin Tablo 2'de pediatristlerin temel başvuru kitabı olan Nelson'da verilen referans aralıklarında da bu tür boşlukların olduğu görülmektedir. Bu boşlukların nasıl doldurulacağı bir tartışma konusudur. Klinik değerlendirmede bu eksiklik test sonucunun hatalı ya da eksik yorumlanmasına neden olabilir. Oysa ki bizim çalışmamızda tanımsız yaş aralıklarının bulunmaması, test sonuçları değerlendirilirken hatalı ya da eksik yorumların oluşmasına engel olacaktır (Tablo 1).

Neonatal bilirubin testine ait bulgular incelendiğinde (Tablo 1), yaş kırılımlarının Tablo 2'de yer alan Nelson kitabına ait yaş kırılımları ile benzerlik taşıdığı görülmektedir. Ancak referans aralığı değerleri karşılaştırıldığında, birinci günden sonrası için farklılıklar göze çarpmaktadır. Bunun olası sebebi birinci günden sonraki sonuçların yenidoğan sarılığı şüphesiyle takip edilen bebeklere ait olmasıdır. Bu durum indirekt referans aralığı belirleme yöntemlerinin genel sorunu olup, daha büyük veri kullanılarak ve teste özgü dışlama kriterleri uygulanarak çözülebilir.

Bu çalışmada en küçük yaş kırılımı 1 gün olarak alınmış olup, her bir güne ait dağılımlar incelenerek, birbirine komşu olan yaş kırılımlarının aynı popülasyona ait olup olmadığı belirlenmiştir. Yüzde 95'in üzerinde benzerlik gösteren kırılımlar birleştirilmiş, altında olanlar ise ayrı bir yaş kırılımı olarak ayrıştırılmıştır. Bu çalışmada, farklı yaş kırılımlarının temsil ettiği popülasyonların benzerliği, literatürde "normal dağılım eğrilerinin örtüşmesi" (overlapping normal distributions) olarak adlandırılan istatistiksel yöntem ile test edilmiştir. Tıp alanında normal dağılım eğrisine sahip iki popülasyonun, örneğin deney ve kontrol grubunun sonuçlarının farklı olup olmadığına karar verebilmek için "T testi" yaygın bir şekilde kullanılmaktadır. T testinde farklılığın anlamlılığı "p değeri"

ile ifade edilirken, bizim çalışmamızda kullandığımız yöntemde ise örtüşmenin (benzerliğin) oranı yüzde olarak ifade edilmektedir. Yani, her iki yöntem de popülasyonlar arasında fark olup olmadığını test etmektedir. Bu çalışmada "normal dağılım eğrilerinin örtüşmesi" yöntemi tercih edilmiştir.[21]

Çalışmamızın olgu grubu yenidoğan evresi olduğu için cinsiyet farklılıkları dikkate alınmadı. Referans aralıkları tek boyutta sadece yaşa göre ayrıştırıldı. Cinsiyet farklılıklarının önemli olduğu durumlarda verilerin kız ve erkekler için ayrı ayrı analiz edilmesi gerekmektedir.

Çalışmanın Kısıtlılıkları

Çalışmamızda hastane veri tabanından elde edilen laboratuvar sonuçlarının hem sağlıklı hem de hasta bireylere ait olması temel bir kısıttır. Bunun nedeni ise hastalık kodlarının ülkemizde hastalığı tanımlamaktan ziyade faturalama ve test istemini yapabilmek için ön koşul olarak kullanılmasıdır.

SONUÇ

Bilgisayar işlemci gücündeki artışlar, yeni özgün yapay zeka ve makine öğrenmesi temelli algoritmalar, yüksek miktarda veri depolamayı sağlayan veri tabanları, referans aralıklarının belirlenmesi için çağdaş bir çözüm geliştirme imkanı sunmaktadır. Bu çalışmada, referans aralıklarının hesaplanmasında temel basamak olan yaş aralıklarını yüksek çözünürlükte belirleyebilen, matematiksel ve istatistiksel temellere dayanan, denetimsiz makine öğrenme algoritması çözümü sunulmuştur. Çalışmada geliştirilen algoritmik yöntemi kullanarak her laboratuvar kendi popülasyon ve analitik yöntemlerine uygun referans aralıklarını kolay, hızlı, güvenli ve ekonomik bir şekilde hesaplayabilecektir.

Etik

Etik Kurul Onayı: Çalışma için Dokuz Eylül Üniversitesi Girişimsel Olmayan Araştırmalar Etik Kurulu'ndan onay alınmıştır (no: 2018/26-22, tarih: 18.10.2018).

Hasta Onayı: Retrospektif çalışmadır.

Hakem Değerlendirmesi: Editörler kurulu ve editörler kurulu dışında olan kişiler tarafından değerlendirilmiştir.

Yazarlık Katkıları

Konsept: O.Y., Ö.A., D.O., S.A.Ö., A.A., S.S., Dizayn: O.Y., Ö.A., S.S., Veri Toplama veya İşleme: O.Y., A.R.Ş., E.Y., M.A.Y., Analiz veya Yorumlama: O.Y., A.R.Ş., D.O., S.A.Ö., A.A., E.Y., M.A.Y., S.S., Literatür Arama: O.Y., Ö.A., A.R.Ş., D.O., S.A.Ö., A.A., E.Y., M.A.Y., S.S., Yazan: O.Y., Ö.A., A.R.Ş., S.S.

Çıkar Çatışması: Yazarlar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemiştir.

Finansal Destek: Çalışmamız için hiçbir kurum ya da kişiden finansal destek alınmamıştır.

Not: Bu makale Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü'nde yürütülen "Algorithmic Learning of Clinically Acceptable Levels of Laboratory Test Results from Electronic Medical Records: Personalized Reference Intervals" başlıklı doktora tez çalışmasından üretilmiştir.

KAYNAKLAR

1. Grasbeck R, Saris NE. Establishment and use of normal values. Scand J Clin Lab Invest. 1969;26:62-3.
2. Solberg HE. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved Recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. J Clin Chem Clin Biochem. 1987;25:337-42.
3. International Federation of Clinical Chemistry (IFCC). Scientific Committee, Clinical Section. Expert Panel on Theory of Reference Values (EPTRV). IFCC Document (1982) stage 2, draft 2, 1983-10-07 with a proposal for an IFCC recommendation. The theory of reference values. Part 2. Selection of individuals for the production of reference values. Clin Chim Acta. 1984;139:205F-13F.
4. Solberg HE, PetitClerc C. Approved recommendation (1988) on the theory of reference values. Part 3. Preparation of individuals and collection of specimens for the production of reference values. Clin Chim Acta. 1988;177:S3-11.
5. Solberg HE, Stamm D. IFCC recommendation: The theory of reference values. Part 4. Control of analytical variation in the production, transfer and application of reference values. J Automat Chem. 1991;13:231-4.
6. Solberg HE. The theory of reference values Part 5. Statistical treatment of collected reference values. Determination of reference limits. J Clin Chem Clin Biochem. 1983;21:749-60.
7. Dybkaer R. International federation of clinical chemistry (IFCC)1),2) the theory of reference values. Part 6. Presentation of observed values related to reference values. J Clin Chem Clin Biochem. 1982;20:841-5.
8. CLSI. Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. CLSI document EP28-A3c. Wayne, PA: Clinical and Laboratory Standards Institute. 2008. Available from: https://clsi.org/media/1421/ep28a3c_sample.pdf
9. Adeli K, Higgins V, Trajcevski K, White-Al Habeeb N. The Canadian laboratory initiative on pediatric reference intervals: A CALIPER white paper. Crit Rev Clin Lab Sci. 2017;54:358-413.
10. Harris EK, Boyd JC. On dividing reference data into subgroups to produce separate reference ranges. Clin Chem. 1990;36:265-70.
11. Lahti A. Partitioning biochemical reference data into subgroups: comparison of existing methods. Clin Chem Lab Med. 2004;42:725-33.
12. Lahti A, Petersen PH, Boyd JC, Rustad P, Laake P, Solberg HE. Partitioning of nongaussian-distributed biochemical reference data into subgroups. Clin Chem. 2004;50:891-900.

13. Mitchell, T. M. Machine Learning. New York: McGraw-Hill. 18;1997.
14. Russell, SJ, Norvig, P. Artificial Intelligence: A Modern Approach (First edition ed.). Prentice Hall. 1995. Available from: http://unina.stidue.net/Intelligenza%20Artificiale/Materiale/Artificial%20Intelligence%20-%20A%20Modern%20Approach%202nd%20ed%20-%20S.%20Russell,%20P.%20Norvig%20(Prentice-Hall,%202003).pdf
15. Arthur, David & Vassilvitskii, Sergei. K-Means++: The Advantages of Careful Seeding. Proc. of the Annu. ACM-SIAM Symp on Discrete Algorithms. 2007;8:1027-35.
16. Reynolds, D. Gaussian Mixture Models. In: Li, S.Z., Jain, A. (eds) Encyclopedia of Biometrics. 2009.
17. Johnson SC. Hierarchical clustering schemes. Psychometrika. 1967;32:241-54.
18. Tukey JW. Exploratory Data Analysis reading, MA: Addison-Wesley. 1977:688. Available from: https://www.stat.berkeley.edu/~brill/Papers/EDASage.pdf
19. Kolmogorov–Smirnov Test. In: The Concise Encyclopedia of Statistics. Springer, New York, NY. 2008. Available from: https://link.springer.com/referenceworkentry/10.1007/978-0-387-32833-1_214
20. Daimon, T. Box-Cox Transformation. In: Lovric, M. (eds) International Encyclopedia of Statistical Science. Heidelberg. 2011. Available from: https://link.springer.com/referenceworkentry/10.1007/978-3-642-04898-2_152
21. Linacre JM. Overlapping normal distributions. Rasch Measurement Transactions, 1996;10:487-8.
22. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17:261-72.
23. Mathematical statistics functions. Available from: https://docs.python.org/3/library/statistics.html
24. Cramér, H. Über eine Eigenschaft der normalen Verteilungsfunktion. Math Z. 1936;41:405-14.
25. International Organization for Standardization. Medical Laboratories-Particular Requirements for Quality and Competence ISO 15189. Geneva: ISO, 2007. Available from: https://www.iso.org/standard/42641.html
26. Kliegman R, Stanton B, St Geme JW, Schor NF, Behrman RE, Nelson WE. Nelson Textbook of Pediatrics. Edition 21 ed. Philadelphia. 2020. Available from: https://www.elsevier.com/books/nelson-textbook-of-pediatrics-2-volume-set/kliegman/978-0-323-52950-1
27. Jones GRD, Haeckel R, Loh TP, et al. Indirect methods for reference interval determination - review and recommendations. Clin Chem Lab Med. 2018;57:20-9.