

RADYOLOJİDE YAPAY ZEKÂ

GÜNCEL UYGULAMALAR
VE KLİNİK YAKLAŞIMLAR

EDITÖR: DR. ÖĞR. ÜYESİ MEHMET ALİ ÖZEL

EĞİTİM
yayınevi

RADYOLOJİDE YAPAY ZEKÂ: GÜNCEL UYGULAMALAR VE KLİNİK YAKLAŞIMLAR

Editör: Dr. Öğr. Üyesi Mehmet Ali Özel

Yayınevi Grubu Genel Başkanı: Yusuf Ziya Aydoğan (yza@egitimyayinevi.com)

Genel Yayın Yönetmeni: Yusuf Yavuz (yusufyavuz@egitimyayinevi.com)

Sayfa Tasarımı: Kübra Konca Nam

Kapak Tasarımı: Eğitim Yayınevi Tasarım Birimi

T.C. Kültür ve Turizm Bakanlığı

Yayıncı Sertifika No: 76780

E-ISBN: 978-625-5014-76-4

1. Baskı, Ekim 2026

Kütüphane Kimlik Kartı

RADYOLOJİDE YAPAY ZEKÂ: GÜNCEL UYGULAMALAR VE KLİNİK YAKLAŞIMLAR

Editör: Dr. Öğr. Üyesi Mehmet Ali Özel

E-ISBN: 978-625-5014-76-4

III+119 s., 160x240 mm

Kaynakça var, dizin yok.

Copyright © Bu kitabın Türkiye'deki her türlü yayın hakkı Eğitim Yayınevi'ne aittir. Bütün hakları saklıdır. Kitabın tamamı veya bir kısmı 5846 sayılı yasanın hükümlerine göre kitabı yayımlayan firmanın ve yazarlarının önceden izni olmadan elektronik/mechanik yolla, fotokopi yoluyla ya da herhangi bir kayıt sistemi ile çoğaltılamaz, yayımlanamaz.

EĞİTİM
yayınevi

Yayınevi Türkiye Ofis:

Konya: Eğitim Yayınevi Tic. Ltd. Şti., Fevzi Çakmak Mah. 10721 Sok. B Blok, No: 16/B, Safakent, Karatay, Konya, Türkiye

İstanbul: Salon Yayınları, Atakent mah., Yasemen sok., No: 4/B, Ümraniye, İstanbul, Türkiye

Santral: +90 332 351 92 85

Editör hatları: +90 533 151 50 42, +90 507 151 50 43

bilgi@egitimyayinevi.com

Yayınevi Amerika Ofis: New York: Egitim Publishing Group, Inc.

P.O. Box 768/Armonk, New York, 10504-0768, United States of America

americaoffice@egitimyayinevi.com

Lojistik ve Sevkiyat Merkezi: Kitapmatik Lojistik ve Sevkiyat Merkezi, Fevzi Çakmak Mah. 10721 Sok. B Blok, No: 16/B, Safakent, Karatay, Konya, Türkiye

İnternet Satış: www.kitapmatik.com.tr

Whatsapp hattı: +90 553 950 50 37

bilgi@kitapmatik.com.tr

Kitabevi Şubesi: Eğitim Kitabevi, Şükran mah. Rampalı 121, Meram, Konya, Türkiye

Whatsapp hattı: +90 501 651 92 85

bilgi@egitimkitabevi.com

EĞİTİM YAYINEVİ
GRUBU

EĞİTİM
yayınevi

SALON
yayınları

Kitapmatik
yayıncılık

kitapmatik
İnternet Kitapçısı

EĞİTİM
kitabevi

İÇİNDEKİLER

YAPAY ZEKANIN TEMELLERİ, ETİK STANDARTLAR VE HUKUKİ SORUMLULUKLAR.....	1
Mehmet Ali Özel	
NÖRORADYOLOJİDE YAPAY ZEKA: İNME, TÜMÖR VE NÖRODEJENERASYON ANALİZİ.....	18
Ömer Önbaş	
TORAKS RADYOLOJİSİNDE YAPAY ZEKA: AKCİĞER NODÜL YÖNETİMİ VE PARANKİMAL HASTALIKLAR.....	34
Büşra Has	
MEME RADYOLOJİSİNDE YAPAY ZEKA: TARAMA VE TANIDA KARAR DESTEK SİSTEMLERİ	48
Derya Güçlü	
ABDOMEN RADYOLOJİSİNDE YAPAY ZEKA: ORGAN SEGMENTASYONU VE LEZYON KARAKTERİZASYONU	59
Mehmet Ali Özel	
KAS-İSKELET SİSTEMİ RADYOLOJİSİNDE YAPAY ZEKA: KIRIK TESPİTİ VE MORFOMETRİK ÖLÇÜMLER.....	71
Elif Nisa Ünlü	
ONKOLOJİK GÖRÜNTÜLEMEDE YAPAY ZEKÂ: RADYOMİK VE TEDAVİ YANITININ ÖNGÖRÜLMESİ.....	85
Ahmet Kerem İmrek	
GİRİŞİMSSEL RADYOLOJİDE YAPAY ZEKA: İŞLEM PLANLAMA VE NAVİGASYON TEKNOLOJİLERİ	103
İbrahim Feyyaz Naldemir	

YAPAY ZEKANIN TEMELLERİ, ETİK STANDARTLAR VE HUKUKİ SORUMLULUKLAR

Mehmet Ali Özel*

GİRİŞ

Radyoloji, sayısal görüntüleme sistemlerinin ve görüntü arşivleme altyapılarının yaygınlaşmasıyla birlikte sağlık hizmetlerinde en fazla veri üreten alanlardan biri haline gelmiştir. Kesitsel görüntüleme yöntemleriyle elde edilen veri miktarındaki artış, hesaplama gücünün gelişmesi ve büyük veri kümelerinin kullanılabilir hale gelmesi, yapay zeka uygulamalarının radyolojide hızla yaygınlaşmasına zemin hazırlamıştır. Özellikle makine öğrenmesi ve derin öğrenme yöntemleri; görüntülerde insan gözüyle ayırt edilmesi güç örüntülerin belirlenmesi, nicel özelliklerin çıkarılması ve tekrarlayan işlemlerin otomatikleştirilmesi açısından önemli olanaklar sunmaktadır (Hosny vd., 2018).

Radyolojide yapay zeka kullanımı yalnızca lezyon tespiti veya görüntü sınıflandırmasıyla sınırlı değildir. Görüntü elde etme ve rekonstrüksiyon, organ ve lezyon segmentasyonu, acil bulguların önceliklendirilmesi, nicel görüntü analizi, raporlama desteği ve kalite güvence gibi iş akışının farklı aşamalarında yapay zekadan yararlanılabilmektedir. Bununla birlikte klinik kullanıma sunulan sistemlerin büyük bölümü, belirli bir görüntüleme yöntemi ve klinik görev için geliştirilmiş dar kapsamlı karar destek araçlarıdır. Bu sistemlerin bağımsız ve genel amaçlı tanı koyan yapılar olarak değerlendirilmesi doğru değildir.

Bir algoritmanın araştırma ortamında yüksek doğruluk veya eğri altında kalan alan değerine ulaşması, klinik uygulamada aynı yararı sağlayacağını tek başına göstermez. Eğitim verisinin niteliği, hasta seçimi, veri ayırım yöntemi, haricî doğrulama, kalibrasyon, farklı hasta gruplarındaki performans ve sistemin mevcut iş akışına nasıl entegre edildiği birlikte değerlendirilmelidir. Retrospektif ve seçilmiş veri kümelerinde başarılı görünen modeller, farklı merkezlerde, cihazlarda veya hasta

*Dr. Öğretim Üyesi, Düzce Üniversitesi, Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, drmaliozel@gmail.com,
ORCID: <https://orcid.org/0000-0001-8817-5769>

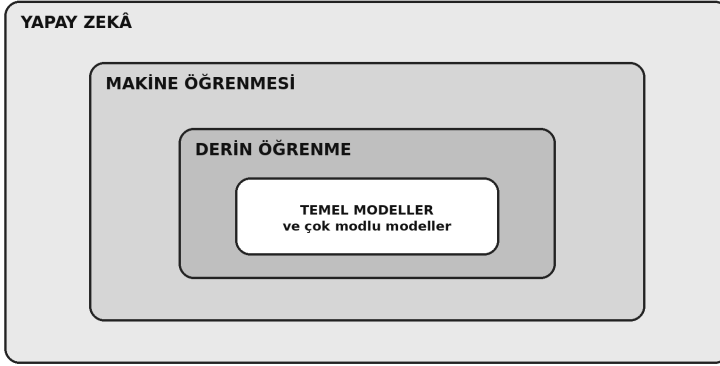
popölasyonlarında performans kaybına uğrayabilir. Bu nedenle deęerlendirmede yalnızca teknik başarıya deęil; klinik yarara, genellenebilirlięe ve uygulama sonrasındaki sürekli performans izlemesine de odaklanılmalıdır (Park ve Han, 2018; Kelly vd., 2019; Tejani vd., 2024).

Yapay zekanın klinik kullanımı hasta mahremiyeti, algoritmik yanlılık, açıklanabilirlik, insan gözetimi ve mesleki sorumluluk gibi etik ve hukuki sorunları da beraberinde getirmektedir. Radyoloęun algoritma çıktısına ne ölçüde güveneceęi, hatalı sonuçlardan hangi paydaşların sorumlu olacağı ve hasta verilerinin hangi koşullarda kullanılabileceęi yalnızca teknik düzenlemelerle çözülebilecek konular deęildir (Geis vd., 2019). Bu bölümde yapay zekanın temel kavramları ve radyolojideki başlıca kullanım alanları açıklanacak; model performansı ve klinik doğrulama ilkeleri, veri güvenlięi, etik standartlar ve algoritmik yanlılık ele alınacaktır. Ayrıca Türkiye, Avrupa Birlięi ve Amerika Birleşik Devletleri'ndeki temel düzenleyici yaklaşımlar incelenerek hekim, saęlık kurumu ve üretici açısından ortaya çıkabilecek sorumluluk alanları tartışılacaktır.

YAPAY ZEKANIN TEKNOLOJİK TEMELLERİ VE KAVRAMSAL ÇERÇEVE

Yapay zeka (YZ), bilgisayar sistemlerinin algılama, örüntü tanıma, dil işleme, tahmin veya karar desteęi gibi normalde insan bilişiyile ilişkilendirilen görevleri yerine getirmesini amaçlayan geniş bir kavramdır. Makine öğrenmesi, sistemin kurallarının tümüyle elle yazılması yerine verideki örüntülerden bir model oluşturmasını saęlayan YZ alt alanıdır. Derin öğrenme ise çok katmanlı yapay sinir aęlarını kullanan makine öğrenmesi yaklaşımıdır. Bu kavramlar birbirine alternatif deęildir; iç içe geçmiş bir hiyerarşı oluşturur (Korkmaz vd., 2024). Temel modeller de bu hiyerarşinin dışında ayrı bir teknoloji deęil, çoęunlukla büyük ölçekli ve öz-denetimli derin öğrenme yöntemleriyle geliştirilen modellerdir (Şekil 1).

Yapay zekâ alanları arasındaki kavramsal ilişki



Denetimli, denetimsiz ve öz-denetimli öğrenme yaklaşımları bu yapıların geliştirilmesinde kullanılabilir.

Şekil 1. Yapay Zeka, Makine Öğrenmesi, Derin Öğrenme ve Temel Modeller Arasındaki İlişki
Öğrenme Yaklaşımları ve Veri Ayrımı

Denetimli öğrenmede model, görüntü ve buna karşılık gelen etiketlerden oluşan örneklerle eğitilir. Etiket; hastalık varlığı, lezyon sınırı, histopatolojik sonuç veya klinik sonlanım olabilir. Denetimsiz öğrenme, önceden tanımlanmış bir hedef etiketi olmadan verideki benzerlikleri ve kümeleri araştırır. Öz-denetimli öğrenmede ise veri içinden türetilen görevler kullanılarak çok sayıda etiketsiz örnekten genel temsil özellikleri öğrenilir. Özellikle temel modellerin geliştirilmesinde öz-denetimli öğrenme önem kazanmıştır (Moor vd., 2023).

Model geliştirme sürecinde eğitim, doğrulama ve test verilerinin işlevleri birbirinden ayrılmalıdır. Eğitim kümesi model parametrelerinin öğrenilmesinde, doğrulama kümesi model ve eşik seçimi gibi geliştirme kararlarında, test kümesi ise geliştirme tamamlandıktan sonra performansın tarafsız biçimde ölçülmesinde kullanılır. Aynı hastaya ait görüntülerin farklı kümelere dağılması veya test verisinin model seçimine dolaylı olarak dahil edilmesi veri sızıntısına ve gerçekte olduğundan daha yüksek performans tahminlerine yol açar.

Geleneksel Makine Öğrenmesi, Radyomiks ve Derin Öğrenme

Geleneksel makine öğrenmesi yaklaşımlarında görüntüden çıkarılacak özellikler çoğunlukla araştırmacı tarafından tanımlanır. Radyomiks; yoğunluk, şekil, tekstür ve dalgacık dönüşümü gibi çok sayıda nicel görüntü özelliğinin çıkarılarak tanısai, prognostik veya prediktif modellerde kullanılmasını sağlayan yöntemler bütünüdür. Bununla birlikte radyomiks

yalnızca elle tasarlanmış özelliklerden ibaret değildir; derin öğrenme ile elde edilen özellikler ve hibrit yaklaşımlar da bu çerçevede değerlendirilebilir (Lambin vd., 2017). Görüntüleme protokolü, rekonstrüksiyon yöntemi ve segmentasyon farklılıkları radyomik özelliklerin tekrarlanabilirliğini etkileyebildiğinden standardizasyon kritik önemdedir.

Derin öğrenmede özellikler büyük ölçüde doğrudan veriden öğrenilir. Evrişimsel sinir ağları, görüntünün yerel ve mekansal özelliklerini işleyebilmesi nedeniyle radyolojik görüntü analizinin gelişiminde merkezi rol oynamıştır. Bununla birlikte güncel sistemler yalnızca evrişimsel ağlarla sınırlı değildir; transformer tabanlı veya evrişimsel ve transformer bileşenlerini birlikte kullanan hibrit mimariler giderek yaygınlaşmaktadır. Bu nedenle tek bir mimarinin tüm radyolojik görevlerde baskın veya üstün olduğu şeklinde bir genelleme yapılmamalıdır.

Temel Modeller, Çok Modlu Sistemler ve Üretken Yapay Zeka

Temel modeller, geniş ve çoğu zaman heterojen veri kümeleri üzerinde önceden eğitilen, daha sonra farklı görevlere uyarlanabilen büyük ölçekli modellerdir. Görüntü, rapor, laboratuvar sonucu ve klinik not gibi birden fazla veri türünü birlikte işleyen çok modlu sistemler; rapor taslağı oluşturma, görüntü-metin eşleştirme, klinik soru yanıtlama ve karar desteği gibi kullanım alanları sunabilir (Moor vd., 2023). Ancak geniş görev kapasitesi, klinik güvenilirlik anlamına gelmez. Üretken modeller gerçeğe aykırı fakat ikna edici içerik üretebilir, belirsizliğini doğru yansıtmayabilir veya eğitim verisindeki yanlılıkları yeniden üretebilir. Bu nedenle çıktıları bağımsız tanı olarak değil, doğrulanması gereken bir öneri veya taslak olarak değerlendirilmelidir (Dünya Sağlık Örgütü [DSÖ], 2024).

RADYOLOJİDE YAPAY ZEKA KULLANIM SENARYOLARI

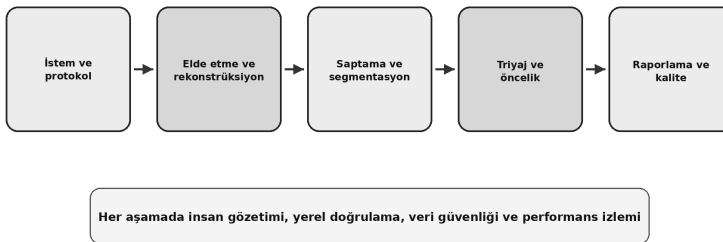
Yapay zeka, radyoloji iş akışının görüntüleme öncesinden sonuçların iletilmesine kadar uzanan farklı aşamalarında kullanılabilir (Şekil 2). İstem bilgilerinin değerlendirilmesi ve uygun protokolün seçilmesi, hasta pozisyonlandırma desteği, çekim parametrelerinin uyarlanması, hareket veya teknik kalite sorunlarının belirlenmesi görüntüleme öncesi ve görüntü elde etme aşamasındaki örneklerdir. BT’de gürültü azaltma ve düşük doz rekonstrüksiyon, MR’de hızlandırılmış görüntüleme ve artefakt düzeltme uygulamaları görüntü kalitesini korurken dozun veya çekim süresinin azaltılmasına katkıda bulunabilir.

Görüntü yorumlama aşamasında bilgisayar destekli saptama ve tanı sistemleri; lezyonların işaretlenmesi, sınıflandırılması ve ölçülmesinde radyoloğa yardımcı olabilir. Segmentasyon sistemleri organ veya lezyon sınırlarını belirleyerek hacim ölçümü, tedavi planlaması ve yanıt değerlendirmesinde standardizasyon sağlayabilir. Bununla birlikte küçük bir Dice skoru farkının klinik olarak anlamlı olup olmadığı, görevin niteliğine ve hatanın anatomik konumuna göre değerlendirilmelidir.

Triyaj uygulamaları intrakraniyal kanama, pulmoner emboli veya pnömotoraks gibi zaman duyarlı bulguları iş listesinde önceliklendirebilir. Ancak triyaj sisteminin yanlış negatif sonuçları, vakaların normal sıralamasını bozarak gecikmeye neden olabileceğinden algoritma çıktısı tek güvenlik katmanı olarak kullanılmamalıdır. Sistemin çalışma listesi, hastane bilgi sistemi ve görüntü arşivleme sistemiyle entegrasyonu; kesinti durumunda izlenecek yol ve uyarıların kime iletileceği önceden tanımlanmalıdır.

Doğal dil işleme ve üretken yapay zeka; rapor taslağı oluşturma, ölçüm ve bulguların yapılandırılması, önceki tetkiklerle karşılaştırma, rapor içi tutarsızlıkların belirlenmesi ve hasta dostu açıklama hazırlanması gibi görevlerde kullanılabilir. Bununla birlikte otomatik oluşturulan raporlar atlanan bulgu, uydurma ifade veya klinik bağlam hatası içerebilir. Nihai raporun radyolog tarafından görüntüler ve klinik bilgilerle karşılaştırılarak onaylanması gerekir. Gerçek yaşam çalışmalarında yapay zekanın bazı süreçlerde zaman kazancı sağlayabildiği bildirilmekle birlikte sonuçlar heterojendir ve her uygulamanın iş yükünü azaltacağı varsayılmamalıdır (Wenderott vd., 2024).

Radyoloji iş akışında yapay zekâ



Şekil 2. Radyoloji İş Akışında Yapay Zeka Kullanım Alanları

PERFORMANS ÖLÇÜTLERİ VE DOĞRULAMA STRATEJİLERİ

Ayrım Gücü, Kalibrasyon ve Klinik Fayda

Bir modelin performansı, amaçlanan klinik göreve uygun birden fazla ölçütle değerlendirilmelidir. Sınıflandırma görevlerinde duyarlılık ve özgüllüğe ek olarak pozitif ve negatif prediktif değerler, doğruluk, F1 skoru ve güven aralıkları verilmelidir. Hastalık prevalansı prediktif değerleri doğrudan etkiler. Dengesiz veri kümelerinde ROC eğrisi altında kalan alan yüksek görünürken pozitif sınıftaki performans yetersiz kalabileceğinden kesinlik-duyarlılık eğrisi de yararlı olabilir. Segmentasyon çalışmalarında Dice katsayısı hacimsel örtüşmeyi, Hausdorff mesafesi ise sınırların en uzak sapmasını değerlendirir; bu ölçütler klinik kabul edilebilirlikle birlikte yorumlanmalıdır (Park ve Han, 2018).

Ayrım gücü, hastalığı olan ve olmayan bireyleri birbirinden ayırabilme kapasitesidir; kalibrasyon ise modelin ürettiği olasılıkların gözlenen riskle ne ölçüde uyumlu olduğunu gösterir. Yüksek AUC'ye sahip bir model kötü kalibre edilmiş olabilir. Bu nedenle kalibrasyon grafikleri ve uygun kalibrasyon ölçütleri raporlanmalıdır. Karar eğrisi analizi gibi yöntemler belirli eşiklerde model kullanımının olası net klinik yararını değerlendirebilir. Bununla birlikte klinik faydanın en güçlü kanıtı, hasta yönetimi, tanı süresi, gereksiz işlem veya klinik sonuçları üzerindeki etkinin prospektif olarak gösterilmesidir.

Doğrulama Tasarımı ve Genellenebilirlik

İç doğrulama, modelin geliştirildiği veri kaynağı içindeki performansını değerlendirir. Eğitim, doğrulama ve test ayrımının hasta düzeyinde yapılması; veri ön işleme, özellik seçimi ve model ayarlama işlemlerinin test kümesinden bağımsız yürütülmesi gerekir. Küçük veri kümelerinde çapraz doğrulama yararlı olabilir; ancak aynı veriyi tekrar tekrar kullanmanın bağımsız bir dış test kümesinin yerini tutmadığı unutulmamalıdır. Model geliştirme ve raporlama sürecinde CLAIM 2024 önerilerinin kullanılması yöntemlerin şeffaflığını ve tekrarlanabilirliğini artırır (Tejani vd., 2024).

Haricî doğrulama, modelin geliştirildiği merkezden bağımsız hasta, cihaz veya protokol verileri üzerinde test edilmesidir. Zamansal doğrulama ise daha sonraki bir dönemde elde edilen verilerle performansın korunup korunmadığını araştırır. Merkezler arasındaki prevalans, hasta seçimi, görüntüleme cihazı, rekonstrüksiyon ve raporlama farklılıkları alan kaymasına neden olabilir. Toplam performansın yanı sıra yaş, cinsiyet, etnik

köken, hastalık şiddeti ve cihaz türü gibi klinik açıdan anlamlı alt gruplar da incelenmelidir.

Klinik kullanımdan önce algoritmanın yerel veriler üzerinde kabul testi yapılmalıdır. İlk aşamada model çıktısının klinisyene gösterilmediği sessiz uygulama, gerçek veri akışındaki teknik sorunların ve beklenmedik hata örüntülerinin belirlenmesini sağlayabilir. Daha sonra kullanıcı davranışı, iş akışı ve güvenlik sonuçlarını değerlendiren erken dönem prospektif çalışmalar planlanabilir. DECIDE-AI, canlı klinik ortamda gerçekleştirilen erken aşama değerlendirmelerinin raporlanması için rehberlik sunar (Vasey vd., 2022).

Tablo 1. Klinik Yapay Zeka Sistemlerinde Doğrulama Düzeyleri

Düzy	Temel soru	Örnek değerlendirme
Teknik iç doğrulama	Model geliştirme verisinde yöntem doğru uygulanmış mı?	Hasta düzeyinde ayırım, çapraz doğrulama, güven aralığı, kalibrasyon
Haricî doğrulama	Model farklı merkez, cihaz ve popülasyonda çalışıyor mu?	Bağımsız merkez verisi, alt grup analizi, hata incelemesi
Zamansal doğrulama	Performans zaman içinde korunuyor mu?	Daha sonraki dönem verisi, protokol ve prevalans değişikliği
Yerel kabul testi	Ürün kurumun gerçek veri akışına uygun mu?	Sessiz uygulama, entegrasyon ve kesinti senaryoları
Klinik etki değerlendirme si	Kullanım hasta yönetimini veya iş akışını iyileştiriyor mu?	Prospektif çalışma, okuyucu çalışması, klinik sonuç ve iş yükü
Kullanım sonrası izlem	Model sahada güvenli ve kararlı kalıyor mu?	Sapma izlemi, olay bildirimi, sürüm kaydı ve yeniden doğrulama

Yaşam Döngüsü ve Kullanım Sonrası İzlem

Yapay zeka sistemi, onay alındıktan veya kuruma kurulduktan sonra değişmeden kalan statik bir ürün olarak görülmemelidir. Cihaz yazılımı, görüntüleme protokolleri, hasta profili ve entegrasyon altyapısı zaman içinde değişebilir. Veri veya kavram kayması model performansında düşüşe yol açabilir. Bu nedenle amaçlanan kullanım, kabul edilen eşikler, izlenecek göstergeler, sorumlu birim, alarm sınırları ve yeniden doğrulama koşulları önceden belirlenmelidir. Performans düşüşü saptandığında sistemin sınırlandırılması veya geçici olarak devre dışı bırakılması da yönetim planının parçası olmalıdır (Cuocolo vd., 2025; Lekadir vd., 2025).

Modelin ve bağılı bileşenlerin sürüm bilgisi, hangi tarihlerde hangi hasta grubunda kullanıldığı ve kullanıcı müdahaleleri kaydedilmelidir. Güncellemeler yalnızca teknik bakım olarak değerlendirilmemeli; kullanım amacını veya klinik performansı etkileyebilecek değişikliklerde yeniden doğrulama yapılmalıdır. Güvenilir yapay zeka yaklaşımı geliştirme, doğrulama, klinik kullanım ve kullanım sonrası izlem basamaklarının tek bir yaşam döngüsü içinde ele alınmasını gerektirir (Şekil 3).



Şekil 3. Klinik Yapay Zeka Sisteminin Yaşam Döngüsü

VERİ YÖNETİMİ, MAHREMİYET VE SİBER GÜVENLİK

Radyolojik görüntüler, raporlar ve bunlarla ilişkilendirilen klinik bilgiler kişisel sağlık verisidir. Türkiye’de sağlık verileri 6698 sayılı Kişisel Verilerin Korunması Kanunu kapsamında özel nitelikli kişisel veri olarak kabul edilir. Verinin sağlık hizmeti, kalite geliştirme, araştırma veya ürün geliştirme amacıyla işlenmesi farklı hukuki dayanaklar gerektirebilir. Bu nedenle her proje için veri sorumlusu ve veri işleyen rolleri, işleme amacı, hukuki sebep, saklama süresi, erişim yetkileri ve aktarım koşulları önceden belirlenmelidir (Kişisel Verileri Koruma Kurumu [KVKK], 2025).

Kimliksizleştirme yalnızca hasta adı ve protokol numarasının DICOM başlığından silinmesi değildir. Özel DICOM alanları, piksele gömülü yazılar, ekran görüntüleri ve serbest metin alanları kimlik bilgisi içerebilir. Kraniyal BT ve MR verilerinden yüzey rekonstrüksiyonu yapılabilmesi de yeniden

tanımlama riski oluşturabilir. Anonimleştirme, verinin kişiyle ilişkisinin geri döndürülemeyecek biçimde ortadan kaldırılmasını; psödönimleştirme ise kimliğin ayrı tutulan bir anahtar yardımıyla yeniden ilişkilendirilebilmesini ifade eder. Psödönimleştirilmiş veri anonim veri sayılmamalıdır (Shahid vd., 2022).

Veri güvenliği için en az ayrıcalık ilkesine dayalı erişim, rol tabanlı yetkilendirme, aktarım ve saklama sırasında şifreleme, erişim kayıtları, düzenli yetki kontrolü, yedekleme ve olay müdahale planı gereklidir. Dış hizmet sağlayıcı veya bulut tabanlı bir sistem kullanıldığında verinin nerede işlendiği, alt yükleniciler, yurt dışı aktarım, silme yükümlülüğü ve güvenlik ihlali bildirimi sözleşmede açıkça düzenlenmelidir. Kişisel sağlık verilerine ilişkin ulusal düzenlemeler 2026 yılında da güncellendiğinden kurum politikalarının yürürlükteki mevzuatla uyumu düzenli olarak gözden geçirilmelidir (Kişisel Sağlık Verileri Hakkında Yönetmelik, 2019/2026).

Federatif öğrenme, veriyi merkez dışına çıkarmadan model güncellemelerinin paylaşılmasını sağlayarak çok merkezli iş birliğini kolaylaştırabilir. Ancak bu yaklaşım tek başına anonimlik veya tam güvenlik sağlamaz; model güncellemelerinden bilgi sızması, merkezler arası veri farklılığı ve güvenli toplulaştırma gibi sorunlar ayrıca yönetilmelidir (Rieke vd., 2020). Benzer biçimde hasta görüntüsü veya raporunun genel amaçlı çevrimiçi üretken yapay zeka hizmetlerine aktarılması, kurum tarafından onaylanmış bir veri işleme altyapısı ve sözleşme bulunmadıkça uygun değildir.

ETİK İLKELER, ALGORİTMİK YANLILIK VE İNSAN GÖZETİMİ

Sağlık alanında yapay zekanın etik değerlendirmesi yalnızca verinin gizliliğine indirgenemez. DSÖ; insan özerkliğinin korunması, insan esenliği ve güvenliğinin geliştirilmesi, şeffaflık ve anlaşılabilirlik, sorumluluk ve hesap verebilirlik, kapsayıcılık ve eşitlik ile sürdürülebilirliği temel ilkeler olarak tanımlamaktadır (DSÖ, 2021). Radyoloji alanındaki çok paydaşlı etik bildiri de yapay zekanın yararı artırırken zararı en aza indirmesi, hak ve özgürlüklere saygı göstermesi ve fayda ile riskleri adil biçimde dağıtması gerektiğini vurgulamaktadır (Geis vd., 2019).

Algoritmik yanlılık, yalnızca eğitim verisinde belirli bir grubun sayıca az olmasından kaynaklanmaz. Hasta seçimi, referans standardın niteliği, etiketleyiciler arasındaki farklılık, görüntüleme cihazı, protokol, hastalık prevalansı ve eksik veriler de sistematik hata oluşturabilir. Bir modelin genel

ortalamada iyi performans göstermesi, bütün alt gruplarda güvenli olduğu anlamına gelmez. Klinik açıdan anlamlı gruplar için duyarlılık, özgüllük, kalibrasyon ve hata türleri ayrı incelenmeli; performans farkları saptandığında veri toplama, yeniden ağırlıklandırma, eşik uyarılama veya kullanım alanını sınırlandırma gibi önlemler değerlendirilmelidir (Koçak vd., 2025).

Şeffaflık; kullanıcıya sistemin amaçlanan kullanımı, eğitim verisinin genel özellikleri, doğrulama koşulları, bilinen sınırlılıkları ve güncel sürümü hakkında yeterli bilgi verilmesini kapsar. Açıklanabilirlik araçları modelin hangi görüntü bölgelerine veya değişkenlere odaklandığını gösterebilir; ancak sunulan açıklamanın doğru veya nedensel olduğu varsayılmamalıdır. Açıklama, güven oluşturmak için kullanılan dekoratif bir çıktı değil, hata incelemesini ve uygun kullanımı destekleyen bir araç olmalıdır (Ahmed vd., 2026).

İnsan gözetimi, algoritma sonucunun ekranda bir hekim tarafından görülmesinden ibaret değildir. Kullanıcının çıktıyı reddedebilmesi, belirsizlik ve başarısızlık durumlarını tanıyabilmesi, gerektiğinde sistemi devre dışı bırakabilmesi ve bu müdahalelerin kayıt altına alınması gerekir. Otomasyon yanlılığı, kullanıcının algoritmanın önerisini kendi değerlendirmesinin önüne koymasına; sistematik olarak algoritmayı görmezden gelmek ise olası yararın kaybına yol açabilir. Eğitimler yalnızca yazılımın nasıl çalıştırılacağını değil, hangi durumlarda güvenilmemesi gerektiğini de kapsamalıdır. Uzun süreli kullanımın mesleki beceri kaybı ve eğitim üzerindeki etkileri izlenmelidir.

Eşitlik ilkesi, yalnızca model performansını değil teknolojiye erişimi de kapsar. Yüksek maliyetli sistemlerin yalnızca belirli merkezlerde kullanılabilmesi, sağlık hizmetindeki mevcut farklılıkları artırabilir. Enerji kullanımı, veri depolama yükü ve gereksiz tekrar işlemler de sürdürülebilirlik bakımından değerlendirilmelidir. Kurum içinde çok disiplinli bir yapay zeka yönetim kurulunun bulunması; radyoloji, bilgi işlem, biyomedikal, hukuk, kalite, veri güvenliği ve hasta güvenliği birimlerinin karar sürecine katılmasını kolaylaştırır. Türk Radyoloji Derneğinin rehberi de ürün seçimi, yerel değerlendirme ve mesleki gözetim açısından ulusal uygulamaya yol gösterir (Türk Radyoloji Derneği [TRD], 2025).

HUKUKİ SORUMLULUK VE DÜZENLEYİCİ ÇERÇEVE

Türkiye’deki Temel Düzenlemeler

Türkiye’de radyolojide kullanılan yapay zeka için bütün sorumluluk alanlarını tek başına düzenleyen bağımsız ve kapsamlı bir mevzuat bulunmamaktadır. Değerlendirme; yazılımın kullanım amacı ve tıbbi cihaz niteliği, sağlık hizmetinin sunumu, hasta hakları, mesleki özen yükümlülüğü, kişisel verilerin korunması, ürün güvenliği ve sözleşme hükümleri birlikte dikkate alınarak yapılır. Üreticisi tarafından tanı, tahmin, prognoz, izlem veya tedavi amacıyla kullanılmak üzere geliştirilen bir yazılım, kullanım amacına bağlı olarak Tıbbi Cihaz Yönetmeliği kapsamına girebilir. Klinik kullanımdan önce ürünün düzenleyici statüsü, uygunluk değerlendirmesi, yetkili kayıtları ve kurumun satın alma koşulları doğrulanmalıdır (Tıbbi Cihaz Yönetmeliği, 2021). Araştırma prototipi ile klinik kullanıma uygun tıbbi cihaz aynı kabul edilmemelidir.

Hasta Hakları Yönetmeliği, hastanın planlanan tıbbi müdahale ve süreç hakkında anlaşılır biçimde bilgilendirilmesini ve rızanın bilgilendirmeye dayanmasını öngörür. Bununla birlikte mevcut mevzuattan her radyolojik incelemede kullanılan bütün yazılım bileşenlerinin hastaya tek tek açıklanması gerektiği biçiminde genel ve koşulsuz bir sonuç çıkarılamaz. Yapay zeka kullanımının hastanın seçimini, işlemin riskini veya klinik kararın niteliğini anlamlı biçimde etkilediği durumlarda bilgilendirmenin kapsamı; kullanım amacı, sistemin rolü, makul alternatifler ve kurumsal politika dikkate alınarak belirlenmelidir (Hasta Hakları Yönetmeliği, 1998/2019). 7557 sayılı Kanun, 2025 yılında bazı tıbbi işlemlerde muvafakatin elektronik ortamda alınabilmesine ilişkin değişiklik yapmış; yapay zeka kullanımına özgü genel bir aydınlatma yükümlülüğü düzenlenmemiştir (7557 Sayılı Kanun, 2025).

Sağlık verisinin işlenmesinde açık rıza her durumda tek veya zorunlu hukuki dayanak değildir. 6698 sayılı Kanun’un güncel 6’ncı maddesinde sağlık hizmeti, tıbbi teşhis ve bakım gibi amaçlar için farklı işleme şartları düzenlenmiştir. Buna karşılık araştırma, ürün geliştirme veya verinin dış hizmet sağlayıcıya aktarılması aynı hukuki sebebe dayanmayabilir. Veri işleme amacı, gereklilik ve ölçülülük ilkeleri ile yurt içi veya yurt dışı aktarım koşulları her kullanım senaryosu için ayrıca değerlendirilmelidir (KVKK, 2025).

Avrupa Birliđi

Avrupa Birliđi'nde tıbbi amaçlı yapay zeka yazılımları açısından Tıbbi Cihaz Tüzüğü (EU) 2017/745 ile Yapay Zeka Tüzüğü (EU) 2024/1689 birlikte önem taşır. Her sađlık yazılımı otomatik olarak yüksek riskli sayılmaz. Bir yapay zeka sisteminin tıbbi cihaz veya cihazın güvenlik bileşeni olması ve ilgili ürünün üçüncü taraf uygunluk deđerlendirmesine tabi bulunması gibi koşullar yüksek risk sınıflandırmasında belirleyicidir. Yüksek riskli sistemlerde risk yönetimi, veri yönetiřimi, teknik dokümantasyon, kayıt tutma, şeffaflık, insan gözetimi, dođruluk, dayanıklılık, siber güvenlik ve kullanım sonrası izlem yükümlölükleri öne çıkar. MDR ile Yapay Zeka Tüzüğü birbirinin alternatifi deđil, uygun ürünlerde eş zamanlı uygulanan tamamlayıcı düzenlemelerdir (Avrupa Birliđi, 2017; Avrupa Birliđi, 2024; Artificial Intelligence Board ve Medical Device Coordination Group, 2025).

Sađlık kuruluřu, sistemi kendi yetkisi altında kullanan taraf olarak yalnızca satın alma aşamasında deđil; kullanım talimatına uygunluk, kullanıcı yetkinliđi, logların saklanması, insan gözetimi ve ciddi olayların bildirilmesi bakımından da sorumluluk taşır. Avrupa Birliđi düzenlemelerinin farklı hükümleri kademeli olarak uygulanmakta ve uygulama takvimi deđişiklik gösterebilmektedir. Bu nedenle baskı öncesinde güncel geçiř hükümleri ve Avrupa Komisyonu rehberleri yeniden kontrol edilmelidir (Kotter vd., 2025).

Amerika Birleşik Devletleri

Amerika Birleşik Devletleri'nde FDA, yapay zeka içeren yazılımı kullanım amacına göre bir tıbbi cihaz yazılım işlevi olarak deđerlendirebilir. Düzenleyici yol, ürünün riskine ve mevcut bir cihaza benzerliđine göre 510(k), De Novo veya Premarket Approval süreçlerinden biri olabilir. FDA'nın yaklaşımı yalnızca piyasaya giriř performansına deđil, toplam ürün yaşam döngüsüne odaklanır. Önceden belirlenmiş deđişiklik kontrol planı, üreticinin gelecekte yapmayı planladığı model deđişikliklerini, bu deđişikliklerin nasıl geliştirileceđini ve dođrulanacađını ve olası etkilerini önceden tanımlamasına imkan verir. Ancak yalnızca yetkilendirilmiş deđişiklikler bu plan kapsamında uygulanabilir (U.S. Food and Drug Administration [FDA], 2025).

Sorumluluğun Paylaşılması

Yapay zeka ile ilişkili bir zararda sorumluluğun her durumda yalnızca radyoloğa, kuruma veya üreticiye ait olduğunu söylemek mümkün değildir. Değerlendirme; hatanın kaynağı, ürünün kullanım amacı, bilinen sınırlılıklar, kullanıcı eğitimi, entegrasyon, klinik bağlam, öngörülebilir yanlış kullanım, kayıtlar ve zararlar nedensellik bağına göre yapılır. Tasarım veya üretim kusuru, yetersiz uyarı ya da güncelleme sorunu üreticiyi; uygunsuz ürün seçimi, yetersiz entegrasyon, eğitim veya izlem kurum yönetimini; klinik bağlamı dikkate almadan kullanma ya da açık bir hatayı göz ardı etme ise kullanıcı hekimi ilgilendirebilir. Aynı olayda birden fazla paydaşın sorumluluğu birlikte gündeme gelebilir (Cestonaro vd., 2023).

Tablo 2. Yapay Zeka Destekli Radyolojide Başlıca Paydaşların Sorumluluk Alanları

Paydaş	Başlıca görevler	Örnek risk alanı
Üretici / geliştirici	Amaçlanan kullanımı tanımlama; güvenlik, klinik performans, teknik dokümantasyon, güncelleme ve kullanım sonrası izlem	Tasarım kusuru, yetersiz doğrulama, hatalı veya eksik uyarı
Sağlık kurumu	Uygun ürün seçimi, yerel kabul testi, entegrasyon, siber güvenlik, eğitim, erişim ve olay yönetimi	Uygunsuz satın alma, entegrasyon hatası, izlem ve eğitim eksikliği
Radyolog / kullanıcı	Amaçlanan kullanım sınırlarına uyma, çıktıyı klinik bağlamda değerlendirme, gerektiğinde reddetme ve önemli hataları bildirme	Otomasyon yanlışlığı, açık hatayı göz ardı etme, kapsam dışı kullanım
Bilgi işlem / biyomedikal	Sürüm ve sistem sürekliliği, veri akışı, loglama, erişim kontrolü ve teknik olayların yönetimi	Yanlış veri eşleşmesi, kesinti, yetkisiz erişim, sürüm karışıklığı
Yönetişim / kalite kurulu	Politika, risk kabulü, performans göstergeleri, alt grup güvenliği ve düzeltici faaliyetlerin gözetimi	Sahipsiz süreç, gecikmiş müdahale, eksik olay bildirimi

Bu nedenle en güçlü hukuki koruma, sorumluluğu tek bir kişiye yükleyen genel ifadelerden değil; görevlerin önceden tanımlanmasından, sistemin sürüm ve kullanım kayıtlarının tutulmasından, yerel doğrulamadan, eğitimden ve hatalara zamanında müdahale edilmesinden doğar. Mevzuat ve uygulama hızla değiştiği için bu bölüm yayımdan hemen önce sağlık hukuku ve tıbbi cihaz mevzuatı konusunda uzman bir hukukçu tarafından son kez gözden geçirilmelidir.

KLİNİK ENTEGRASYON İÇİN ASGARİ KONTROL LİSTESİ

Bir yapay zeka ürününün teknik olarak çalışması, klinik kullanıma hazır olduğu anlamına gelmez. Kuruma özgü karar sürecinde aşağıdaki başlıkların yazılı olarak değerlendirilmesi önerilir.

Tablo 3. Radyolojide Yapay Zeka İçin Klinik Entegrasyon Kontrol Listesi

Kontrol alanı	Sorulacak temel soru	Beklenen kayıt / karar
Amaçlanan kullanım	Sistem hangi modalite, hasta grubu ve klinik görev için geliştirilmiş?	Kapsam ve kapsam dışı kullanım listesi
Düzenleyici durum	Ürün klinik kullanıma uygun tıbbi cihaz statüsünde mi?	Belge, kayıt, sürüm ve kullanım amacı doğrulaması
Kanıt düzeyi	Harici ve prospektif doğrulama var mı; çalışma grubu kuruma benziyor mu?	Kanıt özeti ve kalan belirsizlikler
Yerel performans	Kurum verisinde toplam ve alt grup performansı kabul edilebilir mi?	Yerel kabul testi ve eşik kararı
İş akışı	Çıktı ne zaman, kime ve hangi ekranda gösterilecek?	Süreç şeması, sorumlu roller ve kesinti planı
İnsan gözetimi	Kullanıcı çıktıyı reddedebilir ve hatayı bildirebilir mi?	Yetki, eğitim ve geri bildirim mekanizması
Veri ve siber güvenlik	Veri nerede işleniyor; kim erişiyor; yurt dışına aktarım var mı?	Veri akış haritası, sözleşme ve güvenlik değerlendirmesi
Sürüm ve güncelleme	Güncelleme performansı veya kullanım amacını etkiliyor mu?	Sürüm kaydı ve yeniden doğrulama koşulu
Performans izlemi	Hangi göstergeler, hangi sıklıkta ve kim tarafından izlenecek?	Gösterge, alarm eşiği ve raporlama takvimi
Olay yönetimi	Yanlış sonuç veya sistem kesintisinde nasıl müdahale edilecek?	Olay bildirimi, düzeltici faaliyet ve devre dışı bırakma planı

SONUÇ VE ÖNERİLER

Yapay zeka; görüntü elde etmeden raporlamaya, triyajdan kalite güvenceye kadar radyoloji iş akışının birçok aşamasında yarar sağlayabilir. Ancak araştırma ortamındaki yüksek teknik performans, klinik güvenlik ve hasta yararıyla eş anlamlı değildir. Bir sistemin değeri; amaçlanan kullanıma uygunluğu, bağımsız ve yerel doğrulaması, farklı hasta gruplarındaki performansı, iş akışına etkisi ve kullanım sonrası kararlılığı birlikte değerlendirilerek belirlenmelidir.

Güvenli klinik entegrasyon için sağlık kuruluşlarının çok disiplinli bir yönetim yapısı kurması; ürün seçimi, veri güvenliği, yerel kabul testi, kullanıcı eğitimi, insan gözetimi, sürüm kontrolü ve olay yönetimini yazılı süreçlere bağlaması gerekir. Radyolog, algoritmanın yalnızca nasıl kullanılacağını değil, hangi koşullarda başarısız olabileceğini de bilmelidir. Yapay zeka çıktısı klinik muhakemenin yerine geçmemeli; izlenebilir ve gerektiğinde reddedilebilir bir karar desteği olarak kullanılmalıdır.

Etik ve hukuki sorumluluk, tek bir paydaşa yüklenemeyecek kadar çok katmanlıdır. Üretici güvenli ve doğru tanımlanmış bir ürün geliştirmekten, sağlık kurumu uygun seçme ve entegrasyondan, kullanıcı ise sistemi amaçlanan sınırlar içinde klinik bağlamla birlikte değerlendirmekten sorumludur. Şeffaflık, hesap verebilirlik ve sürekli izlem sağlanmadığında teknik olarak başarılı bir model dahi hasta güvenliği açısından risk oluşturabilir. Geleceğin radyolojisi insan ile makine arasında pasif bir görev paylaşımından çok, sorumlulukları açık ve denetlenebilir bir iş birliği üzerine kurulmalıdır.

KAYNAKÇA

- 7557 Sayılı Kanun. (2025). Sağlıkla İlgili Bazı Kanunlarda ve 663 Sayılı Kanun Hükmünde Karamnamede Değişiklik Yapılmasına Dair Kanun. Resmî Gazete, 24 Temmuz 2025, Sayı: 32965.
- Ahmed, F., Naz, N. S., Khan, S., vd. (2026). Explainable Artificial Intelligence (XAI) in Medical Imaging: A Systematic Review of Techniques, Applications, and Challenges. *BMC Medical Imaging*, 26, 37. DOI: 10.1186/s12880-025-02118-w.
- Artificial Intelligence Board ve Medical Device Coordination Group. (2025). AIB 2025-1/MDCG 2025-6: Interplay Between the Medical Devices Regulation (MDR) & In Vitro Diagnostic Medical Devices Regulation (IVDR) and the Artificial Intelligence Act (AIA). Brussels: European Commission.
- Avrupa Birliği. (2017). Regulation (EU) 2017/745 of the European Parliament and of the Council on Medical Devices. *Official Journal of the European Union*, L 117.
- Avrupa Birliği. (2024). Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence. *Official Journal of the European Union*, L 2024/1689.
- Cestonaro, C., Delicati, A., Marcante, B., Caenazzo, L. ve Tozzo, P. (2023). Defining Medical Liability When Artificial Intelligence Is Applied on Diagnostic Algorithms: A Systematic Review. *Frontiers in Medicine*, 10, 1305756. DOI: 10.3389/fmed.2023.1305756.
- Cuocolo, R., Bernardini, D., Pinto dos Santos, D., vd. (2025). AI Medical Device Post-Market Surveillance Regulations: Consensus Recommendations by the European Society of Radiology. *Insights into Imaging*, 16, 275. DOI: 10.1186/s13244-025-02146-8.
- Dünya Sağlık Örgütü. (2021). Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva: World Health Organization. ISBN: 978-92-4-002920-0.
- Dünya Sağlık Örgütü. (2024). Ethics and Governance of Artificial Intelligence for Health: Guidance on Large Multi-Modal Models. Geneva: World Health Organization. ISBN: 978-92-4-008475-9.
- Geis, J. R., Brady, A. P., Wu, C. C., vd. (2019). Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement. *Radiology*, 293(2), 436-440. DOI: 10.1148/radiol.2019191586.
- Hasta Hakları Yönetmeliği. (1998/2019). Resmî Gazete, 1 Ağustos 1998, Sayı: 23420; değişiklikler 8 Mayıs 2014, Sayı: 28994; 23 Aralık 2016, Sayı: 29927; 16 Ocak 2019, Sayı: 30657.
- Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. ve Aerts, H. J. W. L. (2018). Artificial Intelligence in Radiology. *Nature Reviews Cancer*, 18, 500-510. DOI: 10.1038/s41568-018-0016-5.
- Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G. ve King, D. (2019). Key Challenges for Delivering Clinical Impact with Artificial Intelligence. *BMC Medicine*, 17, 195. DOI: 10.1186/s12916-019-1426-2.
- Kişisel Sağlık Verileri Hakkında Yönetmelik. (2019/2026). Resmî Gazete, 21 Haziran 2019, Sayı: 30808; değişiklik 4 Temmuz 2026, Sayı: 33300.
- Kişisel Verileri Koruma Kurumu. (2025). Özel Nitelikli Kişisel Verilerin İşlenmesine İlişkin Rehber. Ankara: Kişisel Verileri Koruma Kurumu.
- Koçak, B., Ponsiglione, A., Stanzione, A., vd. (2025). Bias in Artificial Intelligence for Medical Imaging: Fundamentals, Detection, Avoidance, Mitigation, Challenges, Ethics, and Prospects. *Diagnostic and Interventional Radiology*, 31(2), 75-88. DOI: 10.4274/dir.2024.242854.
- Korkmaz, M., Sankaya, M. ve Kapukaya, E. E. (2024). Yapay Zeka: Temel Bilgiler ve Önemli Algoritmalar. *TOTBİD Dergisi*, 23, 8-11. DOI: 10.5578/totbid.dergisi.2024.02.
- Kotter, E., D'Antonoli, T. A., Cuocolo, R., vd. (2025). Guiding AI in Radiology: ESR's Recommendations for Effective Implementation of the European AI Act. *Insights into Imaging*, 16, 33. DOI: 10.1186/s13244-025-01905-x.
- Lambin, P., Leijenaar, R. T. H., Deist, T. M., vd. (2017). Radiomics: The Bridge Between Medical Imaging and Personalized Medicine. *Nature Reviews Clinical Oncology*, 14(12), 749-762. DOI: 10.1038/nrclinonc.2017.141.

- Lekadir, K., Frangi, A. F., Porras, A. R., vd. (2025). FUTURE-AI: International Consensus Guideline for Trustworthy and Deployable Artificial Intelligence in Healthcare. *BMJ*, 388, e081554. DOI: 10.1136/bmj-2024-081554.
- Moor, M., Banerjee, O., Abad, Z. S. H., vd. (2023). Foundation Models for Generalist Medical Artificial Intelligence. *Nature*, 616(7956), 259-265. DOI: 10.1038/s41586-023-05881-4.
- Park, S. H. ve Han, K. (2018). Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology*, 286(3), 800-809. DOI: 10.1148/radiol.2017171920.
- Rieke, N., Hancox, J., Li, W., vd. (2020). The Future of Digital Health with Federated Learning. *npj Digital Medicine*, 3, 119. DOI: 10.1038/s41746-020-00323-1.
- Shahid, A., Bazargani, M. H., Banahan, P., vd. (2022). A Two-Stage De-Identification Process for Privacy-Preserving Medical Image Analysis. *Healthcare*, 10(5), 755. DOI: 10.3390/healthcare10050755.
- Tejani, A. S., Klontzas, M. E., Gatti, A. A., vd. (2024). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. *Radiology: Artificial Intelligence*, 6(4), e240300. DOI: 10.1148/ryai.240300.
- Tıbbi Cihaz Yönetmeliği. (2021). *Resmî Gazete*, 2 Haziran 2021, Sayı: 31499 (Mükerrer).
- Türk Radyoloji Derneği. (2025). *Yapay Zeka Uygulamaları Rehberi* (haz. Görüntüleme Bilişimi Çalışma Grubu). Ankara: Türk Radyoloji Derneği. Erişim: <https://www.turkrad.org.tr/assets/DernektenHaberler-Pdf/TRD-YAPAY-ZEKA-UYGULAMALARI-REHBERI.pdf>
- U.S. Food and Drug Administration. (2025). *Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions: Guidance for Industry and Food and Drug Administration Staff*. Silver Spring, MD: FDA.
- Vasey, B., Nagendran, M., Campbell, B., vd. (2022). Reporting Guideline for the Early-Stage Clinical Evaluation of Decision Support Systems Driven by Artificial Intelligence: DECIDE-AI. *Nature Medicine*, 28, 924-933. DOI: 10.1038/s41591-022-01772-9.
- Wenderott, K., Krups, J., Zaruchas, F. ve Weigl, M. (2024). Effects of Artificial Intelligence Implementation on Efficiency in Medical Imaging: A Systematic Literature Review and Meta-Analysis. *npj Digital Medicine*, 7, 265. DOI: 10.1038/s41746-024-01248-9.

NÖRORADYOLOJİDE YAPAY ZEKA: İNME, TÜMÖR VE NÖRODEJENERASYON ANALİZİ

Ömer Önbaş*

GİRİŞ

Yapay zeka (YZ), tıbbi görüntülerden örüntü çıkarma, bulgu saptama, lezyonları sınıflandırma ve nicel ölçüm üretme gibi görevlerde giderek daha fazla araştırılmaktadır. Nöroradyoloji; ince anatomik ayrıntı, çok sekanslı manyetik rezonans (MR) incelemeleri, zamana duyarlı acil durumlar ve seri takip gerektiren hastalıklar nedeniyle bu uygulamaların yoğunlaştığı alanlardan biridir. Bununla birlikte bir algoritmanın belirli bir veri kümesinde yüksek performans göstermesi, farklı merkezlerde aynı doğrulukla çalışacağı veya hasta sonuçlarını iyileştireceği anlamına gelmez. YZ sistemleri bu nedenle radyoloğun yerine geçen otonom tanı araçları olarak değil, uygun biçimde doğrulandığında görüntü yorumunu ve klinik iş akışını destekleyen yazılımlar olarak değerlendirilmelidir (Erickson et al., 2017; Pesapane et al., 2018; Wagner et al., 2023).

Nöroradyolojide YZ çalışmalarının önemli bir bölümü akut iskemik inme, intrakraniyal kanama ve beyin tümörleri üzerinde yoğunlaşmıştır. Bu alanlarda amaç; kritik incelemeleri iş listesinde öne almak, büyük damar tıkanıklığını işaretlemek, enfarktüs çekirdeği ve hipoperfüze dokuyu nicelleştirmek, kanama veya metastaz odaklarını saptamak ve tümör hacmini izlemektir. Beyaz cevher hastalıkları, multipl skleroz (MS) ve nörodejeneratif hastalıklarda ise otomatik segmentasyon, beyin volumetrisi ve hastalık ilerlemesinin öngörülmesi öne çıkmaktadır. Ancak akut inme triyayı ile hafif bilişsel bozukluktan demansa ilerleme öngörüsü aynı klinik olgunluk düzeyinde değildir; her uygulama kendi hedefi, veri yapısı ve hata maliyeti içinde değerlendirilmelidir (Kelly et al., 2022; Wagner et al., 2023).

Bu bölümde nöroradyolojide başlıca YZ görevleri tanımlanacak; akut iskemik inme, intrakraniyal kanama, beyin tümörleri, beyaz cevher hastalıkları, MS ve nörodejeneratif hastalıklardaki uygulamalar klinik

* Prof. Dr., Düzce Üniversitesi Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, Türkiye. E-posta: omeronbas@duzce.edu.tr. ORCID: <https://orcid.org/0000-0003-1174-0437>

kanıtlarıyla birlikte ele alınacaktır. Ayrıca dış doğrulama, iş akışı entegrasyonu, insan gözetimi ve düzenleyici izlem gibi klinik kullanımı belirleyen kısıtlılıklar tartışılacaktır.

NÖRORADYOLOJİDE YAPAY ZEKA GÖREVLERİ VE PERFORMANSIN DEĞERLENDİRİLMESİ

Temel Görevler

Nöroradyolojik YZ uygulamaları beş temel görev altında incelenebilir. Saptama, görüntüde belirli bir bulgunun veya anatomik hedefin varlığını ve konumunu işaretler; büyük damar tıkanıklığı, intrakraniyal kanama veya küçük metastaz odağı buna örnektir. Sınıflandırma, incelemeyi ya da lezyonu önceden tanımlanmış kategorilerden birine atar. Segmentasyon, lezyonun veya anatomik yapının piksel ya da voksel düzeyinde sınırlarını çıkarır. Nicelleştirme, segmentasyon sonucunu hacim, oran, lezyon sayısı veya atrofi ölçüsü gibi tekrarlanabilir çıktılara dönüştürür. Öngörü modelleri ise görüntüleme verisini klinik, laboratuvar veya demografik değişkenlerle birleştirerek tedavi yanıtı ya da hastalık ilerlemesi gibi gelecekteki sonuçları tahmin etmeye çalışır (Erickson et al., 2017; Wagner et al., 2023).

Evrişimli sinir ağları, görüntünün yerel uzamsal özelliklerini öğrenebilmesi nedeniyle sınıflandırma ve saptama görevlerinde yaygın olarak kullanılmaktadır. U-Net ve türevleri ise kodlayıcı-çözücü yapıları sayesinde segmentasyon çalışmalarında öne çıkar. Derin öğrenmeyi belirli bir katman sayısı ile tanımlamak doğru değildir; derinlik, ağın mimarisine ve probleme göre değişir. Benzer biçimde bu ağların insan beyninin yapısını doğrudan taklit ettiği söylenemez. Klinik açıdan önemli olan, mimarinin adı değil; hedeflenen görev, eğitim verisinin temsil gücü, doğrulama tasarımı ve modelin hata örüntüsüdür.

Performans Ölçütlerinin Yorumlanması

Saptama ve sınıflandırma modellerinde duyarlılık, özgüllük, pozitif ve negatif öngörü değerleri, doğruluk ve ROC eğrisi altında kalan alan (AUC) sık kullanılan ölçütlerdir. Duyarlılık ve özgüllük eşik seçimine bağlıdır; pozitif ve negatif öngörü değerleri ise hastalık prevalansından etkilenir. Bu nedenle yalnızca AUC veya doğruluk bildirilmesi klinik yararı göstermeye yetmez. Özellikle dengesiz veri kümelerinde yüksek doğruluk, az görülen fakat kritik bir bulgunun kaçırılmasını gizleyebilir. Performansın güven

aralıklarıyla, hasta düzeyinde ayrılmış bir test kümesinde ve mümkünse bağımsız dış veri üzerinde sunulması gerekir (Kelly et al., 2022).

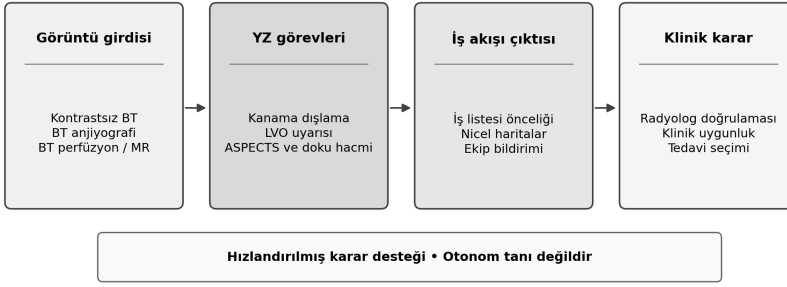
Segmentasyon performansında Dice benzerlik katsayısı, modelin ürettiği maske ile referans maske arasındaki örtüşmeyi 0 ile 1 arasında özetler. Dice değerinin yüksek olması klinik açıdan önemli küçük lezyonların eksiksiz saptandığını tek başına garanti etmez; küçük lezyonlarda hacme duyarlı ölçütler kararsız olabilir. Bu nedenle lezyon bazlı duyarlılık, yanlış pozitif sayısı, hacim farkı ve sınır uzaklığı gibi tamamlayıcı ölçütler de değerlendirilmelidir. Referans standardın birden fazla uzman tarafından oluşturulması ve uzmanlar arası değişkenliğin bildirilmesi, model performansının gerçekçi yorumlanmasını sağlar.

AKUT İSKEMİK İNME

Klinik Sorun ve Görüntüleme İş Akışı

Akut iskemik inmede görüntüleme; intrakraniyal kanamayı dışlamak, büyük damar tıkanıklığını belirlemek, geri dönüşümsüz hasar görmüş dokuyu ve risk altındaki ancak kurtarılabilir dokuyu değerlendirmek için kullanılır. Mekanik trombektominin yararı, uygun görüntüleme ve klinik ölçütlerle seçilmiş hastalarda DAWN çalışmasında son sağlıklı görülme zamanından 6–24 saat, DEFUSE 3 çalışmasında ise 6–16 saat aralığında gösterilmiştir. Bu veriler tüm inme tedavilerinin zaman penceresinin koşulsuz olarak 24 saate uzadığı biçiminde yorumlanmamalıdır; geç pencerede tedavi seçimi hasta, damar ve doku özelliklerine bağlıdır (Albers et al., 2018; Nogueira et al., 2018).

YZ, bu iş akışının farklı aşamalarında görüntülerin önceliklendirilmesi, büyük damar tıkanıklığı uyarısı, otomatik ASPECTS hesaplanması ve perfüzyon haritalarının nicelleştirilmesi için kullanılabilir. En olgun kullanım biçimi, algoritmanın şüpheli incelemeyi iş listesinde öne alması ve ilgili ekibe bildirim üretmesidir. Nihai karar; kaynak görüntüleri, klinik tabloyu, semptom başlangıç zamanını, kollateral dolaşımı ve tedaviye ilişkin kontrendikasyonları birlikte değerlendiren hekim tarafından verilmelidir (Murray et al., 2020; Kim et al., 2023).



Şekil 1. Akut iskemik inmede yapay zeka destekli görüntüleme iş akışının şematik gösterimi.

Büyük Damar Tıkanıklığı ve Triyaj

BT anjiyografi üzerinde geliştirilen modeller internal karotis arterin intrakraniyal segmentleri ile orta serebral arterin proksimal dallarındaki tıkanıklıkları saptamayı ve şüpheli incelemeler için otomatik uyarı oluşturmayı hedefler. Bu yaklaşım özellikle yoğun acil servislerde raporlama sırasını değiştirebilir ve görüntünün doğru ekibe ulaşmasına kadar geçen süreyi azaltabilir. Bununla birlikte tıkanıklığın distal yerleşimli olması, damar hareketi, yetersiz kontrast bolusu, venöz kontaminasyon, metal artefaktı ve kronik stenozlar yanlış sonuçlara yol açabilir. Literatürde bildirilen yüksek tanısal performansa rağmen çalışmalar arasında hasta seçimi, hedef damar tanımı ve doğrulama yöntemi açısından belirgin heterojenlik bulunmaktadır (Murray et al., 2020).

Bir triyaj aracının klinik değeri yalnızca tıkanıklık saptama duyarlılığıyla ölçülmemelidir. Bildirimin doğru kişiye ulaşması, yanlış uyarı yükü, iş listesindeki gerçek zaman kazancı, radyolog doğrulama süresi ve kapı–tedavi aralığı üzerindeki etki birlikte incelenmelidir. Uyarı sisteminin kesintiye uğraması veya yanlış negatif sonuç üretmesi olasılığı nedeniyle standart inme iş akışı algoritmaya bağımlı hale getirilmemelidir.

Otomatik ASPECTS

ASPECTS, anterior dolaşım iskemisinde kontrastsız BT üzerindeki erken iskemik değişiklikleri topografik olarak puanlar. Otomatik ASPECTS yazılımları hemisferik asimetri, dansite farklılıkları ve bölgesel özelliklerden yararlanarak tekrarlanabilir bir skor üretmeyi amaçlar. Meta-analizlerde yapay zeka temelli değerlendirme ile uzman puanlaması arasında anlamlı uyum bildirilmiş olmakla birlikte performans; enfarktüsün yerleşimine,

görüntü kalitesine, rekonstrüksiyon parametrelerine ve referans okuyuculara göre değişmektedir (Adamou et al., 2023).

Otomatik skorun gözlemciler arası değişkenliği tamamen ortadan kaldırdığı söylenemez. Bazal gangliyon düzeyi, kronik enfarktılar, lökoaraiyozis, hareket ve ışın sertleşmesi artefaktları skoru etkileyebilir. Bu nedenle yazılımın ürettiği toplam puanla birlikte bölgesel işaretlemeler kaynak kesitler üzerinden kontrol edilmeli; skor klinik kararın tek belirleyicisi olarak kullanılmamalıdır.

Perfüzyon Görüntüleme ve Doku Nicelleştirmesi

BT perfüzyon ve MR perfüzyon incelemelerinde amaç, geri dönüşümsüz hasar görmüş enfarktüs çekirdeği ile ağır hipoperfüzyona rağmen kurtarılabilir olabilecek dokuyu yaklaşık olarak ayırmaktır. Otomatik yazılımlar hareket düzeltme, arteriyel giriş fonksiyonu seçimi, dekonvolüsyon ve eşikleme basamaklarını bir araya getirerek hacimsel haritalar üretir. DEFUSE 3'te kullanılan RAPID yaklaşımında enfarktüs çekirdeği için normal dokuya göre rölatif serebral kan akımının %30'un altında olması, hipoperfüze doku için ise Tmax gecikmesinin 6 saniyeden uzun olması kullanılmıştır (Albers et al., 2018).

Bu değerler evrensel biyolojik sınırlar değildir. Sonuçlar kullanılan yazılıma, cihaz ve protokole, hareket artefaktına, kardiyak debiye, proksimal darlıklara ve gecikme düzeltme yöntemine göre değişebilir. Küçük çekirdek hacimlerinde birkaç mililitrelik ölçüm farkı hasta sınıflamasını etkileyebilir; erken reperfüzyon veya çok kötü kollateral dolaşım da otomatik haritaların klinik tabloyla uyumsuz görünmesine neden olabilir. Haritalar kaynak görüntülerle birlikte incelenmeli ve hatalı segmentasyon olasılığı göz önünde bulundurulmalıdır.

İNTRAKRANİYAL KANAMA

Saptama, Sınıflandırma ve Hacim Ölçümü

Kontrastsız beyin BT, akut intrakraniyal kanamanın değerlendirilmesinde temel görüntüleme yöntemidir. YZ sistemleri inceleme düzeyinde kanama varlığını belirlemek, epidural, subdural, subaraknoid, intraparenkimal ve intraventriküler alt tipleri sınıflandırmak, kanama hacmini segmentasyonla ölçmek ve pozitif olabilecek incelemeleri önceliklendirmek amacıyla geliştirilmiştir. Bu kullanım, özellikle yüksek iş yükünde raporlama gecikmesini azaltabilecek bir karar desteği sunar; ancak kanamanın

yokluğunu kesinleştiren bağımsız bir dışlama testi olarak değerlendirilmemelidir.

RSNA-ASNR 2019 Beyin BT Kanama Yarışması için üç kurumdan yaklaşık 874 bin görüntü toplanmış ve altmıştan fazla nöroradyoloğun katkısıyla etiketlenmiştir. Bu veri kümesi, çok sayıda modelin ortak bir test ortamında karşılaştırılmasını sağlamıştır (Flanders et al., 2020). Yarışmada birinci olan ve daha sonra yayımlanan Wang ve arkadaşlarının modeli, RSNA test verisinde kesit düzeyinde 0,950 duyarlılık ve 0,944 özgüllük; inceleme düzeyinde ise 0,945 duyarlılık ve 0,995 özgüllük bildirmiştir. Bu değerler veri kümesini tanımlayan Flanders çalışmasına değil, ilgili algoritma çalışmasına aittir (Wang et al., 2021).

Kanama hacmi ve yayılımının otomatik segmentasyonu seri takip, büyüme değerlendirmesi ve araştırma amaçlı nicelleştirme açısından yararlı olabilir. Bununla birlikte küçük subaraknoid kanamalar, kafa tabanına komşu hiperdensiteler, kalsifikasyon, kontrast madde, hareket ve metal artefaktları önemli hata kaynaklarıdır. Vacek ve arkadaşları farklı YZ yazılımlarının kanama sınırlandırmasında değişken sonuçlar verebildiğini göstermiştir; güncel tek merkezli çalışmalar ise doğru olgularda zaman kazancı bildirse de gerçek dünya performansının yerel prevalans ve iş akışından etkilendiğini ortaya koymaktadır (Vacek et al., 2024; Damar et al., 2026).

Mikrokanama Kavramı

BT tabanlı akut kanama saptaması ile serebral mikrokanama değerlendirmesi birbirinden ayrılmalıdır. Serebral mikrokanamalar çoğunlukla hemosiderin birikimine bağlı küçük duyarlılık odaklarıdır ve gradient eko T2* ya da duyarlılık ağırlıklı MR görüntülerinde değerlendirilir. Kontrastsız BT'de çok küçük mikrokanamaların rutin olarak saptandığını ileri sürmek doğru değildir. BT algoritmaları için daha uygun hedefler akut hiperdens kanama, alt tip sınıflandırması, kitle etkisi ve hacim ölçümüdür.

BEYİN TÜMÖRLERİ

Metastaz Saptama ve Tümör Sınıflandırması

Beyin metastazları, özellikle çok sayıda küçük lezyonun bulunduğu hastalarda radyolog için zaman alıcı bir saptama görevidir. Otomatik saptama modelleri kontrastlı T1-ağırlıklı MR görüntülerinde aday odakları işaretleyerek ikinci okuyucu gibi çalışabilir. İki ayrı meta-analizde metastaz

saptama algoritmaları için yaklaşık %89–90 düzeyinde havuzlanmış duyarlılık bildirilmiştir. Bununla birlikte lezyon çapı küçüldükçe duyarlılık azalmakta; damar kesitleri, dural yapılar ve artefaktlar yanlış pozitif odaklara yol açabilmektedir (Cho et al., 2021; Ozkara et al., 2023).

Tümör sınıflandırma modelleri gliom, meningiom, hipofiz tümörü veya normal inceleme gibi sınıfları ayırmayı hedefleyebilir. Bu modellerin önemli bir bölümü herkese açık, seçilmiş ve iki boyutlu görüntü kümeleri üzerinde geliştirilmiştir. Asif ve arkadaşları InceptionV3 ile Xception mimarilerini birleştiren bir topluluk modelinde %98,30 doğrulama doğruluğu bildirmiştir; bu değer duyarlılık değildir. Çalışmanın 2.870 özgün görüntüsünün veri artırma yöntemleriyle 10.000 görüntüye çıkarılmış olması, modelin bağımsız ve çok merkezli klinik performansını göstermez. Veri artırılmış görüntülerin eğitim ve test kümelerine hasta düzeyinde ayrılmaması halinde veri sızıntısı riski de doğabilir. Bu nedenle yüksek doğrulama doğruluğu, patolojik tanının yerini alabilecek bir “sanal biyopsi” olarak yorumlanmamalıdır (Asif et al., 2025).

Radyomik ve radyogenomik yaklaşımlar; tümör derecesi, moleküler alt tip veya tedavi yanıtıyla ilişkili olabilecek nicel görüntü özelliklerini araştırmaktadır. Bu alanın klinik potansiyeli yüksek olsa da MR cihazı, sekans, yeniden örnekleme, segmentasyon ve özellik çıkarma adımlarındaki farklılıklar sonuçların tekrarlanabilirliğini sınırlayabilir. Patoloji gerektiren bir hastalıkta görüntüleme temelli sınıflandırma, uygun doğrulama yapılmadan kesin histolojik veya moleküler tanı olarak sunulmamalıdır (Campion & Iv, 2023).

Tümör Segmentasyonu ve Nicelleştirme

Beyin tümörlerinde otomatik segmentasyon; kontrastlanan tümör, nekroz, tümör çekirdeği ve T2/FLAIR hiperintens infiltratif alan gibi alt bölgelerin hacmini ölçmeyi amaçlar. Bu çıktılar cerrahi ve radyoterapi planlamasını destekleyebilir, seri incelemelerde hacim değişimini nicelleştirebilir ve klinik araştırmalarda daha tekrarlanabilir sonlanım ölçütleri sağlayabilir. Ancak kistik değişiklik, kanama, postoperatif kavite, tedaviye bağlı kontrastlanma ve infiltratif sınırlar referans maske oluşturmayı güçleştirir.

Huang ve arkadaşlarının BraTS verileriyle yürüttüğü çalışmada yalnızca kontrastlı T1-ağırlıklı ve FLAIR sekanslarının kullanıldığı model, test kümesinde kontrastlanan tümör için 0,867, tümör çekirdeği için 0,926 Dice değerine ulaşmıştır. Sonuç, azaltılmış sekans kombinasyonlarının belirli veri kümelerinde yeterli olabileceğini düşündürmektedir; ancak dört sekansın

linik olarak gereksiz olduğunu kanıtlamaz. Çok merkezli dış doğrulama, farklı tümör tipleri ve tedavi sonrası incelemeler üzerindeki performans gösterilmeden gerçek dünya entegrasyonuna ilişkin güçlü sonuç çıkarılmamalıdır (Huang et al., 2025).

Tedavi Sonrası Değerlendirme

Tedavi sonrası gliom görüntülemesinde gerçek progresyon, psödoprogresyon, radyasyon nekrozu ve anti-anjiyojenik tedaviye bağlı psödoyanıt ayrımı güçtür. Konvansiyonel MR bulgularına perfüzyon, difüzyon, spektroskopi veya PET verilerinin eklenmesi tanısal değerlendirmeyi destekleyebilir. YZ modelleri bu çok boyutlu veriyi birleştirerek progresyon olasılığı üretmeyi amaçlar; ancak referans standardın histopatoloji yerine klinik ve görüntüleme takibine dayanması etiket belirsizliği yaratabilir (Kessler & Bhatt, 2018).

Sun ve arkadaşları 77 glioblastom hastasında kontrastlı T1-ağırlıklı görüntülerden elde edilen radyomik özelliklerle psödoprogresyonu gerçek progresyondan ayırmayı araştırmış; model için %78 duyarlılık ve %61 özgüllük bildirmiştir. Sonuçlar araştırma açısından umut verici olmakla birlikte sınırlı örneklem, tek merkezli tasarım ve orta düzey özgüllük nedeniyle klinik kararın tek başına modele bırakılmasını desteklemez (Sun et al., 2021). İnfiltratif hücre yoğunluğunu gösterdiği ileri sürülen haritalar veya başka görevler, aynı çalışmayla ilişkilendirilmemeli ve ayrı kanıtlarla desteklenmelidir.

BEYAZ CEVHER HASTALIKLARI VE MULTİPL SKLEROZ

Beyaz Cevher Hiperintensitelerinin Nicelleştirilmesi

Beyaz cevher hiperintensiteleri T2-ağırlıklı ve FLAIR görüntülerde yüksek sinyalli alanlar olarak izlenir; yaşlanma ve serebral küçük damar hastalığı başta olmak üzere demiyelinizan, inflamatuvar, toksik, metabolik ve genetik birçok süreçte görülebilir. Bu nedenle tüm beyaz cevher lezyonlarını tek bir hastalık biyobelirteci olarak değerlendirmek doğru değildir. STRIVE yaklaşımı lezyonların tanımlanması ve raporlanması için standart terimler önermiş, güncel STRIVE-2 çalışması ise dağılım ve çoklu görüntüleme belirteçlerinin önemini vurgulamıştır (Duering et al., 2023).

Otomatik segmentasyon, görsel Fazekas derecelendirmesini tamamlayan sürekli bir hacim ölçüsü sağlayabilir. Dadar ve arkadaşlarının farklı sınıflandırma yöntemlerini karşılaştırdığı çalışma, algoritmaların beyaz cevher hiperintensitelerini nicelleştirebildiğini ancak yöntem ve veri

kümesine göre performansın değiştiğini göstermiştir (Dadar et al., 2017). Klinik yorumda toplam hacmin yanı sıra periventriküler ve derin yerleşim, lakünler, mikrokranamalar, kortikal atrofi ve hastanın vasküler risk profili birlikte ele alınmalıdır. BT veya PET verisinin otomatik olarak eklenmesi her klinik sorunda yarar sağlamaz; multimodal yaklaşım ancak hastalığa özgü bir gerekçe ve doğrulanmış ek katkı varsa kullanılmalıdır.

Multipl Skleroz

MS'de YZ için en doğrudan nöroradyolojik hedefler; T2/FLAIR hiperintens lezyonların segmentasyonu, yeni veya genişleyen lezyonların seri incelemelerde saptanması, kontrastlanan lezyonların belirlenmesi ve beyin ile omurilik atrofisinin nicelleştirilmesidir. Otomatik lezyon hacmi, görsel değerlendirmeyi destekleyebilir; ancak lezyon sayısı ve yükü klinik özüllülüğün tüm boyutlarını açıklamaz. Juxtakortikal ve infratentoryal küçük lezyonlar, damar boşlukları, hareket, kısmi hacim etkisi ve farklı FLAIR protokolleri hata kaynaklarıdır.

MSSEG yarışması, ortak bir veri altyapısı üzerinde farklı segmentasyon yöntemlerini karşılaştırmış ve uzmanlar arası değişkenliği görünür kılmıştır (Commowick et al., 2018). McKinley ve arkadaşları, tek merkezde eğitilen modellerin merkez dışı veride performans kaybı yaşayabildiğini; ek anatomik etiketleme ve çoklu yapı segmentasyonunun bazı mimarilerde genellenebilirliği artırabildiğini göstermiştir (McKinley et al., 2021). Bu bulgu, yalnızca yüksek iç doğrulama Dice değerlerine dayalı klinik kullanımın sakıncasını açıkça ortaya koymaktadır.

Görüntüleme dışındaki giyilebilir sensörler, yürüme analizi ve dijital bilişsel testler MS izleminde araştırılabilir; ancak bunlar nöroradyolojik görüntü analiziyle aynı başlık altında ayrıntılandırılmamalıdır. Dijital ikizler veya hedefe yönelik nanorobotik ilaç taşıma sistemleri ise rutin klinik nöroradyoloji uygulaması değildir. Bu bölümün odağı, klinik uygulanabilirliği daha yüksek olan lezyon saptama, segmentasyon, atrofi ölçümü ve görüntüleme–klinik veri bütünleştirilmesi olmalıdır. Storelli ve arkadaşlarının çalışması, MR ve klinik verilerle MS hastalarında nörokognitif kötüleşme öngörüsünü inceleyen bu tür bir çok değişkenli yaklaşıma örnektir; sonuçlar MS bağlamında yorumlanmalıdır (Storelli et al., 2026).

NÖRODEJENERATİF HASTALIKLAR

Yapısal MR ve Otomatik Volumetri

Nörodejeneratif hastalıklarda YZ uygulamalarının temel amacı, görsel değerlendirmeyi nicel beyin ölçümleriyle desteklemek ve farklı klinik tablolar arasındaki örüntüleri araştırmaktır. T1-ağırlıklı üç boyutlu MR verilerinden toplam beyin hacmi, kortikal kalınlık, hipokampal ve medial temporal hacimler, ventrikül hacmi ve bölgesel atrofi örüntüleri çıkarılabilir. Bu ölçümler yaş ve kafa büyüklüğüne göre düzeltilerek normatif veri tabanlarıyla karşılaştırılabilir. Ancak segmentasyon hataları, kafa içi hacim düzeltilmesi, tarayıcı ve sekans değişikliği ile normatif örneklemin demografik yapısı sonuçları belirgin biçimde etkiler.

Alzheimer hastalığında medial temporal atrofi önemli bir görüntüleme bulgusudur; ancak yaşlanma, vasküler hastalık, epilepsi ve başka nörodejeneratif süreçlerde de görülebilir. Otomatik volumetri, klinik öykü ve nöropsikolojik değerlendirmeden bağımsız bir tanı testi değildir. Basaia ve arkadaşları tek bir yapısal MR incelemesini kullanan derin öğrenme modeliyle Alzheimer hastalığı, stabil hafif bilişsel bozukluk ve ilerleyici hafif bilişsel bozukluk gruplarını ayırmayı araştırmıştır. Çalışma, yapısal MR'dan hastalıkla ilişkili örüntülerin öğrenilebileceğini gösterse de, seçilmiş araştırma kohortlarındaki performansın heterojen rutin hasta popülasyonuna doğrudan aktarılması gerekir (Basaia et al., 2019).

Hafif Bilişsel Bozukluktan Demansa İlerleme

Hafif bilişsel bozukluk klinik olarak heterojen bir durumdur; hastaların bir bölümü Alzheimer demansına ilerlerken bir bölümü stabil kalabilir veya bilişsel olarak düzelebilir. Bu nedenle ilerleme öngörüsü, yalnızca mevcut tanıyı sınıflandırmaktan daha zor ve klinik olarak daha anlamlı bir hedeftir. Spasov ve arkadaşları yapısal MR verisini yaş, cinsiyet ve bilişsel ölçümlerle birleştiren parametre açısından verimli bir derin öğrenme yaklaşımıyla ilerleme riskini tahmin etmeyi araştırmıştır (Spasov et al., 2019). Bu tür modellerin zaman ufku, ilerleme tanımı, kayıp takip oranı ve tedavi etkileri açıkça belirtilmelidir.

“Beyin yaşı”, sağlıklı bireylerin MR görüntülerinden kronolojik yaşı öğrenen bir modelin belirli bir hastadaki tahmini ile gerçek yaş arasındaki farkı ifade eder. Derin öğrenme temelli beyin yaşı ölçümlerinin araştırma kohortlarında güvenilir ve kalıtsal bir özellik gösterebildiği bildirilmiştir (Cole et al., 2017). Bununla birlikte beyin yaşı hastalığa özgü değildir;

vasküler hasar, gelişimsel farklılıklar, yaşam tarzı ve teknik etkenler aynı ölçümü etkileyebilir. Bu nedenle bireysel hastada “hızlanmış yaşlanma” tanısı koymak veya tedavi başlatmak için tek başına kullanılmamalıdır.

Multimodal Modeller ve Açıklanabilirlik

Yapısal MR; FDG-PET, amiloid veya tau PET, bilişsel testler, genetik ve klinik verilerle birleştirildiğinde tek bir modalitenin yakalayamadığı bilgileri sağlayabilir. Qiu ve arkadaşlarının çok merkezli çalışması, multimodal derin öğrenmenin Alzheimer demansı değerlendirmesinde görüntüleme ve klinik veriyi bir araya getirebildiğini göstermiştir (Qiu et al., 2022). Bununla birlikte her yeni veri türünün maliyet, erişilebilirlik ve eksik veri sorunu vardır. Multimodal model, yalnızca tek modaliteli modele göre dış doğrulanmış ve klinik açıdan anlamlı bir ek yarar sağlıyorsa tercih edilmelidir.

Açıklanabilir YZ yöntemleri modelin önem verdiği bölgeleri ısı haritaları veya özellik katkı değerleriyle gösterebilir. Bu çıktılar modelin teknik denetimini kolaylaştırır da modelin doğru patofizyolojik gerekçeyle karar verdiğini kanıtlamaz. Isı haritasının hipokampusu vurgulaması, sınıflandırmanın nedensel olarak hipokampal atrofiye dayandığını tek başına göstermez. Açıklanabilirlik çıktıları; hata analizi, veri sızıntısı kontrolü, dış doğrulama ve uzman değerlendirmesinin yerine geçmemelidir.

Frontotemporal demans, Lewy cisimcikli demans ve parkinsoniyen sendromlarda da sınıflandırma ve örüntü tanıma modelleri araştırılmaktadır. Ancak hastalık sıklığı, patolojik doğrulamanın azlığı ve üst üste binen görüntüleme örüntüleri nedeniyle bu alanlarda klinik genellenebilirlik daha sınırlıdır. Bölümün ana klinik mesajı, YZ'nin nörodejeneratif hastalıkları görünür bir tek işaretle tanıyan bir sistem olmadığı; çok kaynaklı veriyi nicel ve tekrarlanabilir biçimde birleştirmeyi amaçlayan bir karar desteği olduğudur.

KLİNİK KANIT, ENTEGRASYON VE KISITLILIKLAR

Kanıt Düzeyinin Değerlendirilmesi

Bir YZ çalışmasının kanıt değeri, yalnızca bildirilen performans ölçütüyle belirlenmez. Eğitim, doğrulama ve test kümelerinin hasta düzeyinde ayrılması; aynı hastaya ait seri görüntülerin farklı kümelere dağılmaması; referans standardın tanımlanması; eksik verilerin yönetimi ve önceden belirlenmiş analiz planı temel gerekliliklerdir. Tek merkezli retrospektif test, geliştirmenin ilk aşaması için yararlı olabilir; klinik kullanıma geçiş için

bağımsız çok merkezli dış doğrulama ve ardından prospektif iş akışı değerlendirmesi gerekir (Kelly et al., 2022).

Araştırma sonuçları da birbirinden ayrılmalıdır. Algoritmanın AUC veya Dice değerindeki artış teknik performansı; raporlama süresindeki azalma iş akışı etkisini; tedaviye kadar geçen sürenin veya fonksiyonel sonlanımın iyileşmesi ise klinik etkiyi gösterir. İlk iki düzeydeki başarı, üçüncü düzeyi otomatik olarak kanıtlamaz. Özellikle akut inme ve kanama triyajında zaman kazancı bildirilirken yanlış uyarı yükü, uyarı yorgunluğu ve gerçek hasta sonuçları da izlenmelidir.

Etki Alanı Değişimi ve Genellenebilirlik

Modelin geliştirildiği veri dağılımı ile uygulandığı klinik ortam arasındaki fark etki alanı değişimi olarak adlandırılır. Cihaz üreticisi, manyetik alan gücü, kesit kalınlığı, rekonstrüksiyon çekirdeği, kontrast zamanlaması, hasta yaşı ve hastalık prevalansı bu değişime katkıda bulunabilir. Model yeni merkezde sessiz biçimde performans kaybedebilir. MS segmentasyonunda merkez dışı veride görülen performans düşüşü, nöroradyolojide bu riskin somut bir örneğidir (McKinley et al., 2021).

Yerel kullanımdan önce temsili bir veri kümesinde sessiz mod doğrulama yapılmalı; duyarlılık, yanlış pozitif yükü, kalibrasyon ve başarısız işlem oranı izlenmelidir. Yazılım veya görüntüleme protokolü güncellendiğinde doğrulama tekrarlanmalıdır. Düzenleyici onay, bir ürünün tanımlanmış kullanım amacı için belirli gereklilikleri karşıladığını gösterir; her kurumda aynı klinik etkiyi sağlayacağını veya her yeni yazılım sürümünün değişmeden çalışacağını garanti etmez.

İnsan Gözetimi ve İş Akışı Tasarımı

YZ çıktısı, radyoloğun kaynak görüntülere erişimini engellemeden ve belirsizliği görünür kılacak biçimde PACS ya da iş listesine entegre edilmelidir. Kritik uyarının hangi kullanıcıya, ne kadar sürede ve hangi yedek kanalla iletileceği önceden tanımlanmalıdır. Kullanıcı yalnızca modelin işaretlediği bölgelere odaklanırsa işaretlenmeyen bulgular gözden kaçabilir; bu durum otomasyon yanlışlığının önemli bir biçimidir. Radyolog, model sonucundan bağımsız sistematik inceleme alışkanlığını korumalıdır.

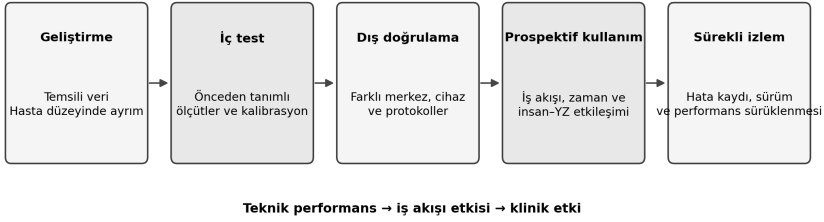
Başarısızlık durumu da tasarımın parçasıdır. Eksik seri, bozuk DICOM başlığı, desteklenmeyen protokol veya sunucu kesintisinde sistem sessizce sonuç üretmemeli; hatayı görünür biçimde bildirmelidir. Kurumlar yanlış negatif ve yanlış pozitif olayları kaydetmeli, klinik önemi yüksek hatalar için

düzenli kalite toplantıları yapmalı ve model sürümlerini izlemelidir. YZ'nin güvenli kullanımı tek bir satın alma kararı değil, sürekli kalite güvencesi sürecidir.

Tablo 1. Nöroradyolojide başlıca yapay zeka uygulamalarının kanıt ve klinik olgunluk açısından özeti

Klinik alan	Girdi ve görev	Başlıca çıktı	Kanıt/olgunluk	Temel kısıt
Akut inme	BTA'da LVO saptama ve triyaj	Uyarı, iş listesi önceliği	Klinik kullanıma en yakın alanlardan biri	Distal tıkanıklık, bolus ve artefakt; yanlış uyarı yükü
Akut inme	Kontrastsız BT'de ASPECTS	Bölgesel skor ve işaretleme	Karar desteği; yazılıma göre değişken	Kronik lezyon, lökoaraiyozis, erken değişiklikler
Akut inme	BT/MR perfüzyonda segmentasyon	Çekirdek ve hipoperfüze doku hacmi	Seçilmiş geç pencere hastalarında yerleşik nicelleştirme	Yazılım ve protokol bağımlı eşikler
İntrakraniyal kanama	Kontrastsız BT'de saptama, alt tip ve hacim	Triyaj ve nicel izlem	Olgunlaşan karar desteği	Küçük SAH, kafa tabanı, kalsifikasyon ve metal
Beyin metastazı	Kontrastlı MR'da odak saptama	İkinci okuyucu ve aday işaretleme	Umut verici; klinik doğrulama gerekli	Küçük lezyonlarda düşük duyarlılık ve yanlış pozitifler
Gliom	Segmentasyon ve tedavi sonrası sınıflandırma	Alt bölge hacmi, progresyon olasılığı	Segmentasyon olgunlaşıyor; yanıt öngörüsü araştırma düzeyinde	Etiket belirsizliği ve merkezler arası protokol farkı
Beyaz cevher / MS	FLAIR lezyon segmentasyonu ve seri değişim	Yeni/büyüyen lezyon yükü	Araştırma ve seçilmiş klinik kullanım	Etki alanı değişimi; küçük lezyonlar ve damar boşlukları
Nöro-dejenerasyon	Volumetri, sınıflandırma ve ilerleme öngörüsü	Atrofi ölçüsü ve risk olasılığı	Volumetri destekleyici; prognoz modelleri araştırma düzeyinde	Normatif veri, hastalık heterojenliği ve özgüllük sorunu

Not: Klinik olgunluk, düzenleyici onay anlamına gelmez. Her ürün tanımlanmış kullanım amacı, sürümü ve yerel doğrulama sonuçları içinde değerlendirilmelidir. BTA: BT anjiyografi; LVO: büyük damar tıkanıklığı; SAH: subaraknoid kanama; MS: multipl skleroz.



Şekil 2. Nöroradyolojide bir yapay zeka aracının klinik kullanıma geçiş basamakları.

SONUÇ VE ÖNERİLER

Nöroradyolojide YZ; akut inme ve intrakraniyal kanamada triyaj, beyin tümörlerinde saptama ve segmentasyon, beyaz cevher hastalıklarında lezyon nicelleştirmesi ve nörodejeneratif hastalıklarda volumetri ile risk modellemesi açısından önemli bir karar desteği potansiyeline sahiptir. En yüksek klinik uygunluk, açık biçimde tanımlanmış ve zaman açısından kritik görevlerde görülmektedir. Buna karşılık tümör moleküler sınıflandırması, psödoprogresyon ayrımı ve demansa ilerleme öngörüsü gibi karmaşık hedeflerde kanıtlar çoğunlukla araştırma düzeyindedir.

Klinik uygulamada yüksek teknik performansın ötesine geçilmelidir. Modelin bağımsız ve çok merkezli veride doğrulanması, yerel iş akışında gerçek zamanlı test edilmesi, hasta düzeyindeki etkisinin araştırılması ve yazılım yaşam döngüsü boyunca izlenmesi gerekir. İnsan gözetimi, başarısızlıkların görünür kılınması ve otomasyon yanlışlığının azaltılması güvenli uygulamanın temel koşullarıdır. YZ'nin gerçek değeri radyoloğun yerini almasında değil; tekrarlayan görevleri standartlaştırması, nicel bilgi üretmesi ve kritik bulguların daha erken değerlendirilmesine katkıda bulunmasında yatmaktadır.

KAYNAKÇA

- Adamou, A., Beltsios, E. T., Bania, A., Gkana, A., Kastrup, A., Chatziioannou, A., Politi, M., & Papanagioutou, P. (2023). Artificial intelligence-driven ASPECTS for the detection of early stroke changes in non-contrast CT: A systematic review and meta-analysis. *Journal of NeuroInterventional Surgery*, 15(e2), e298–e304. <https://doi.org/10.1136/jnis-2022-019447>
- Albers, G. W., Marks, M. P., Kemp, S., Christensen, S., Tsai, J. P., Ortega-Gutierrez, S., ... Lansberg, M. G. (2018). Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *The New England Journal of Medicine*, 378(8), 708–718. <https://doi.org/10.1056/NEJMoal713973>
- Asif, R. N., Naseem, M. T., Ahmad, M., Mazhar, T., Khan, M. A., Khan, M. A., Al-Rasheed, A., & Hamam, H. (2025). Brain tumor detection empowered with ensemble deep learning approaches from MRI scan images. *Scientific Reports*, 15, 15002. <https://doi.org/10.1038/s41598-025-99576-7>
- Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., ... Alzheimer's Disease Neuroimaging Initiative. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical*, 21, 101645. <https://doi.org/10.1016/j.nicl.2018.101645>
- Campion, A., & Iv, M. (2023). Brain tumor imaging: Review of conventional and advanced techniques. *Seminars in Neurology*, 43(6), 867–888. <https://doi.org/10.1055/s-0043-1776765>
- Cho, S. J., Sunwoo, L., Baik, S. H., Bae, Y. J., Choi, B. S., & Kim, J. H. (2021). Brain metastasis detection using machine learning: A systematic review and meta-analysis. *Neuro-Oncology*, 23(2), 214–225. <https://doi.org/10.1093/neuonc/noaa232>
- Cole, J. H., Poudel, R. P. K., Tsagkrasoulis, D., Caan, M. W. A., Steves, C., Spector, T. D., & Montana, G. (2017). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163, 115–124. <https://doi.org/10.1016/j.neuroimage.2017.07.059>
- Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., ... Barillot, C. (2018). Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. *Scientific Reports*, 8, 13650. <https://doi.org/10.1038/s41598-018-31911-7>
- Dadar, M., Maranzano, J., Misquitta, K., Anor, C. J., Fonov, V. S., Tartaglia, M. C., ... Alzheimer's Disease Neuroimaging Initiative. (2017). Performance comparison of 10 different classification techniques in segmenting white matter hyperintensities in aging. *NeuroImage*, 157, 233–249. <https://doi.org/10.1016/j.neuroimage.2017.06.009>
- Damar, Ö., Yaman, M., & Gokdemir, M. T. (2026). Diagnostic accuracy and time efficiency of artificial intelligence for intracranial hemorrhage detection on emergency brain CT. *Diagnostics*, 16(14), 2219. <https://doi.org/10.3390/diagnostics16142219>
- Duering, M., Biessels, G. J., Brodtmann, A., Chen, C., Cordonnier, C., de Leeuw, F. E., ... Wardlaw, J. M. (2023). Neuroimaging standards for research into small vessel disease—advances since 2013. *The Lancet Neurology*, 22(7), 602–618. [https://doi.org/10.1016/S1474-4422\(23\)00131-X](https://doi.org/10.1016/S1474-4422(23)00131-X)
- Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. *Radiographics*, 37(2), 505–515. <https://doi.org/10.1148/rq.2017160130>
- Flanders, A. E., Prevedello, L. M., Shih, G., Halabi, S. S., Kalpathy-Cramer, J., Ball, R., ... RSNA-ASNR 2019 Brain Hemorrhage CT Annotators. (2020). Construction of a machine learning dataset through collaboration: The RSNA 2019 Brain CT Hemorrhage Challenge. *Radiology: Artificial Intelligence*, 2(3), e190211. <https://doi.org/10.1148/ryai.2020190211>
- Huang, J., Yagmurlu, B., Molleti, P., Lee, R., Vanderploeg, A., Noor, H., Bareja, R., Li, Y., Iv, M., & Itakura, H. (2025). Brain tumor segmentation using deep learning: High performance with minimized MRI data. *Frontiers in Radiology*, 5, 1616293. <https://doi.org/10.3389/fradi.2025.1616293>
- Kelly, B. S., Judge, C., Bollard, S. M., Clifford, S. M., Healy, G. M., Aziz, A., ... Killeen, R. P. (2022). Radiology artificial intelligence: A systematic review and evaluation of methods (RAISE). *European Radiology*, 32(11), 7998–8007. <https://doi.org/10.1007/s00330-022-08784-6>

- Kessler, A. T., & Bhatt, A. A. (2018). Brain tumour post-treatment imaging and treatment-related complications. *Insights into Imaging*, 9(6), 1057–1075. <https://doi.org/10.1007/s13244-018-0661-y>
- Kim, M., Jung, S. C., Kim, S. C., Kim, B. J., Seo, W. K., & Kim, B. (2023). Proposed protocols for artificial intelligence imaging database in acute stroke imaging. *Neurointervention*, 18(3), 149–158. <https://doi.org/10.5469/neuroint.2023.00339>
- McKinley, R., Wepfer, R., Aschwanden, F., Grunder, L., Muri, R., Rummel, C., ... Wiest, R. (2021). Simultaneous lesion and brain segmentation in multiple sclerosis using deep neural networks. *Scientific Reports*, 11, 1087. <https://doi.org/10.1038/s41598-020-79925-4>
- Murray, N. M., Unberath, M., Hager, G. D., & Hui, F. K. (2020). Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: A systematic review. *Journal of NeuroInterventional Surgery*, 12(2), 156–164. <https://doi.org/10.1136/neurintsurg-2019-015135>
- Nogueira, R. G., Jadhav, A. P., Haussen, D. C., Bonafe, A., Budzik, R. F., Bhuva, P., ... Jovin, T. G. (2018). Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *The New England Journal of Medicine*, 378(1), 11–21. <https://doi.org/10.1056/NEJMoal706442>
- Ozkara, B. B., Chen, M. M., Federau, C., Karabacak, M., Briere, T. M., Li, J., & Wintermark, M. (2023). Deep learning for detecting brain metastases on MRI: A systematic review and meta-analysis. *Cancers*, 15(2), 334. <https://doi.org/10.3390/cancers15020334>
- Pesapane, F., Codari, M., & Sardanelli, F. (2018). Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *European Radiology Experimental*, 2, 35. <https://doi.org/10.1186/s41747-018-0061-6>
- Qiu, S., Miller, M. I., Joshi, P. S., Lee, J. C., Xue, C., Ni, Y., ... Kolachalama, V. B. (2022). Multimodal deep learning for Alzheimer's disease dementia assessment. *Nature Communications*, 13, 3404. <https://doi.org/10.1038/s41467-022-31037-5>
- Spasov, S., Passamonti, L., Duggento, A., Liò, P., Toschi, N., & Alzheimer's Disease Neuroimaging Initiative. (2019). A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *NeuroImage*, 189, 276–287. <https://doi.org/10.1016/j.neuroimage.2019.01.031>
- Storelli, L., Mistri, D., Mastropasqua, A., Grosselle, M., Preziosa, P., Rossi, L., Filippi, M., & Rocca, M. A. (2026). Explainable artificial intelligence to predict neurocognitive disorder progression in multiple sclerosis using MRI and clinical data. *European Journal of Neurology*, 33(5), e70568. <https://doi.org/10.1111/ene.70568>
- Sun, Y. Z., Yan, L. F., Han, Y., Nan, H. Y., Xiao, G., Tian, Q., ... Cui, G. B. (2021). Differentiation of pseudoprogression from true progression in glioblastoma patients after standard treatment: A machine learning strategy combined with radiomics features from T1-weighted contrast-enhanced imaging. *BMC Medical Imaging*, 21, 17. <https://doi.org/10.1186/s12880-020-00545-5>
- Vacek, A., Mair, G., White, P., Bath, P. M., Muir, K. W., Al-Shahi Salman, R., ... Wardlaw, J. M. (2024). Evaluating artificial intelligence software for delineating hemorrhage extent on CT brain imaging in stroke. *Journal of Stroke and Cerebrovascular Diseases*, 33(1), 107512. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2023.107512>
- Wagner, D. T., Tilmans, L., Peng, K., Niedermeier, M., Rohl, M., Ryan, S., ... Nguyen, X. V. (2023). Artificial intelligence in neuroradiology: A review of current topics and competition challenges. *Diagnostics*, 13(16), 2670. <https://doi.org/10.3390/diagnostics13162670>
- Wang, X., Shen, T., Yang, S., Lan, J., Xu, Y., Wang, M., Han, X., Zhang, J., & Zhang, X. (2021). A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages in head CT scans. *NeuroImage: Clinical*, 32, 102785. <https://doi.org/10.1016/j.nicl.2021.102785>

TORAKS RADYOLOJİSİNDE YAPAY ZEKA: AKCİĞER NODÜL YÖNETİMİ VE PARANKİMAL HASTALIKLAR

Büşra Has*

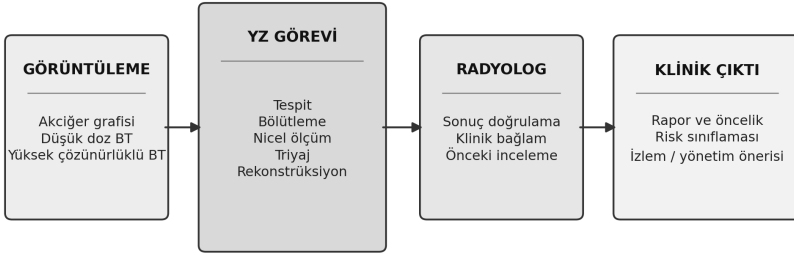
GİRİŞ

Toraks radyolojisi; yüksek inceleme hacmi, geniş hastalık spektrumu ve çok sayıda görüntünün kısa sürede değerlendirilmesini gerektirmesi nedeniyle yapay zeka (YZ) uygulamalarının en yoğun geliştirildiği alanlardan biridir. Akciğer grafisinde anormallik tespiti, bilgisayarlı tomografide (BT) pulmoner nodül analizi, diffüz parankimal hastalıkların nicel değerlendirilmesi, acil incelemelerin önceliklendirilmesi ve düşük doz görüntülerin rekonstrüksiyonu başlıca uygulama alanlarını oluşturmaktadır. Bununla birlikte, bir algoritmanın belirli bir veri kümesinde yüksek ayırt etme performansı göstermesi, farklı hasta gruplarında aynı sonucu vereceği veya hasta sonuçlarını iyileştireceği anlamına gelmez (Hosny vd., 2018; Lee vd., 2019; Hwang vd., 2025).

Toraks görüntülemesinde YZ sistemleri çoğunlukla dar ve önceden tanımlanmış görevler için geliştirilir. Bir model yalnızca anormallik varlığını sınıflandırabilir, lezyonu işaretleyebilir, hacim ölçebilir, iş listesini yeniden sıralayabilir veya görüntü gürültüsünü azaltabilir. Bu görevlerin her biri farklı referans standartları, performans ölçütleri ve klinik riskler taşır. Bu nedenle tespit, bölütleme, karakterizasyon, prognoz ve triyaj sonuçları birbirinin yerine kullanılmamalıdır.

Bu bölümde toraks radyolojisindeki başlıca YZ uygulamaları, klinik kullanım amacı, kanıt düzeyi ve sınırlılıklarıyla birlikte ele alınmaktadır. Temel yaklaşım, YZ çıktısının bağımsız bir tanı olarak değil; radyolog değerlendirmesi, klinik bilgiler, önceki incelemeler ve uygun yönetim kılavuzlarıyla bütünleştirilen bir karar desteği olarak kullanılmasıdır.

*Dr. Öğretim Üyesi, Düzce Üniversitesi, Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, hasbusra@gmail.com, ORCID: <https://orcid.org/0000-0003-2963-2696>



YZ çıktısı nihai karar değildir; klinik sorumluluk radyolog ve ilgili klinik ekibindedir.

Şekil 1. Toraks görüntülemesinde yapay zeka destekli klinik iş akışı.

AKCİĞER GRAFİSİNDE YAPAY ZEKA

Anormallik Tespiti ve Okuyucu Desteği

Akciğer grafisi yaygın, erişilebilir ve düşük maliyetli bir yöntem olmakla birlikte; anatomik süperpozisyon, taşınabilir çekimlerdeki teknik farklılıklar ve çok çeşitli bulgular nedeniyle yorumlanması güç olabilir. Özellikle küçük nodüller, apikal veya hiler lezyonlar, retrokardiyak opasiteler ve küçük pnömotoraksler algısal hatalara açıktır. Derin öğrenme tabanlı sistemler; nodül veya kitle, konsolidasyon, plevral efüzyon, pnömotoraks, atelettazi ve kardiyomegali gibi bulguların tespitinde ikinci okuyucu olarak değerlendirilmektedir (Hwang vd., 2019; Nam vd., 2021; Bennani vd., 2023).

Okuyucu çalışmalarında YZ desteği bazı bulgular için radyolog duyarlılığını artırabilmektedir; ancak katkının büyüklüğü okuyucunun deneyimine, hastalık prevalansına, seçilen eşik değerine ve olgu karmaşıklığına göre değişir. Bir modelin zenginleştirilmiş bir test kümesinde radyolog düzeyinde sonuç vermesi, ardışık gerçek yaşam olgularında veya farklı cihaz ve popülasyonlarda aynı performansın korunacağını göstermez. Ayrıca yanlış pozitif işaretlemeler gereksiz BT istemine ve ek iş yüküne, yanlış negatif sonuçlar ise otomasyon yanlılığına yol açabilir (Hwang vd., 2025).

YZ desteğinin okuma süresine etkisi de tüm incelemelerde aynı değildir. Normal olarak değerlendirilen grafilere güveni artırarak süreyi kısaltabildiği, buna karşılık modelin anormallik işaretlediği incelemelerde ek doğrulama gereksinimi nedeniyle belirgin zaman kazancı sağlamayabildiği bildirilmiştir (Shin vd., 2023). Bu nedenle verimlilik değerlendirmesinde

yalnızca ortalama okuma süresi değil; yanlış uyarı yükü, ek inceleme oranı, rapor değişikliği ve kritik bulgunun sonuçlanma süresi birlikte incelenmelidir.

Klinik Kullanım Sınırı

Akciğer grafisindeki YZ çıktıları tek başına normal veya anormal kararı için yeterli kabul edilmemelidir. Taşınabilir anteroposterior grafiler, yoğun bakım hastaları, tüp ve kateterler, postoperatif değişiklikler ve yetersiz inspirasyon gibi durumlar model performansını etkileyebilir. Kuruma alınacak bir sistemin yerel veriyle sessiz biçimde test edilmesi, farklı alt gruplarda kalibrasyonunun değerlendirilmesi ve yanlış negatif olgular için güvenli iş akışı oluşturulması gerekir.

PULMONER NODÜL YÖNETİMİNDE YAPAY ZEKA

Klinik Bağlam ve Yönetim Sistemleri

Pulmoner nodüller tarama BT'lerinde veya başka bir nedenle gerçekleştirilen toraks BT incelemelerinde insidental olarak saptanabilir. Bu iki klinik bağlam aynı yönetim sistemiyle ele alınmaz. Lung-RADS, akciğer kanseri tarama programlarındaki incelemelerin sınıflandırılması ve izlem önerilerinin standardizasyonu için geliştirilmiştir. Fleischner Topluluğu önerileri ise 35 yaş ve üzerindeki erişkinlerde insidental nodüllere yöneliktir; tarama hastalarında, immünsüprese bireylerde ve bilinen primer malignitesi bulunan hastalarda kullanılmaz (MacMahon vd., 2017; Christensen vd., 2024).

Düşük doz BT ile taramanın yüksek riskli popülasyonlarda akciğer kanserine bağlı mortaliteyi azaltabildiği gösterilmiştir; ancak tarama incelemelerinde saptanan nodüllerin büyük bölümü benigndir (Aberle vd., 2011). YZ, nodül tespiti ve ölçümünü destekleyebilir; buna karşılık daha fazla benign nodülün işaretlenmesi geri çağırma oranını ve izlem yükünü artırabilir. Klinik değer, yalnızca kaç nodül bulunduğuyla değil; kanser saptama oranı, yanlış pozitif yönetimi ve gereksiz girişimlerin dengesiyle değerlendirilmelidir (Hwang vd., 2025).

Nodül Tespiti

Klasik bilgisayar destekli tespit sistemleri, yüksek duyarlılıkla çok sayıda aday lezyon belirleyip sonraki aşamada damar kesitleri, skar alanları ve plevral yapılardan kaynaklanan yanlış pozitifleri azaltmaya çalışır. Güncel üç boyutlu derin öğrenme modelleri, ham BT verisinden daha karmaşık

uzamsal özellikleri öğrenilmekte ve tespit performansını geliştirebilmektedir. LIDC-IDRI ve LUNA16 gibi açık veri kümeleri yöntemlerin karşılaştırılmasına önemli katkı sağlamıştır (Armato vd., 2011; Setio vd., 2017; Min vd., 2022).

Bununla birlikte, ölçütlerin yorumlanması klinik bağlamdan bağımsız yapılmamalıdır. Duyarlılık, inceleme başına yanlış pozitif sayısı ile birlikte bildirilmelidir. Küçük, subsolid, plevraya veya damara komşu nodüller; kalın kesit, hareket artefaktı ve farklı rekonstrüksiyon algoritmaları tespit performansını düşürebilir. Ayrıca ölçüt veri kümelerinde seçilmiş ve ayrıntılı biçimde anotasyonlanmış nodüller üzerinde elde edilen başarı, ardışık klinik popülasyonda aynı sonucu garanti etmez. Rutin klinik popülasyonlarda yapılan dış değerlendirmeler bu nedenle daha yüksek kanıt değeri taşır (Murchison vd., 2022).

Hua ve arkadaşlarının çalışması, BT'de nodüllerin derin öğrenme ile benign-malign sınıflandırılmasına yönelik erken örneklerden biridir; çalışma nodül tespitinden çok sınıflandırma görevine odaklanmaktadır (Hua vd., 2015). Tespit ve karakterizasyon görevlerinin ayrı tutulması, modelin klinik kullanım amacının doğru anlaşılması açısından önemlidir.

Bölütleme ve Hacim Ölçümü

Nodülün güvenilir biçimde bölütlenmesi; çap, hacim, solid bileşen ve zaman içindeki değişimin hesaplanması için temel basamaktır. Manuel ölçümler özellikle küçük ve subsolid nodüllerde gözlemciler arası değişkenlik gösterebilir. U-Net ve türevleri gibi üç boyutlu mimariler otomatik bölütleme sağlayabilse de sonuçlar nodülün kenar özelliği, komşu damar veya plevra ilişkisi, kesit kalınlığı ve rekonstrüksiyon parametrelerinden etkilenir (Chen vd., 2022; Ma vd., 2020).

Otomatik hacim ölçümünde küçük mutlak farklar, küçük nodüllerde oransal olarak büyük değişimlere dönüşebilir. Bu nedenle seri incelemelerde benzer çekim ve rekonstrüksiyon protokolleri kullanılmalı; modelin bölütleme sınırı radyolog tarafından kontrol edilmeli ve büyüme kararı yalnızca tek bir otomatik ölçüme dayandırılmamalıdır.

Malignite Riskinin Değerlendirilmesi

Malignite riski; nodül boyutu, yoğunluğu, kenar özellikleri, lokalizasyonu ve büyüme davranışının yanı sıra yaş, sigara öyküsü ve önceki malignite gibi klinik değişkenlerle ilişkilidir. Derin öğrenme modelleri görüntüyü uçtan uca analiz edebilir; radyomik yaklaşımlar ise şekil, yoğunluk ve doku

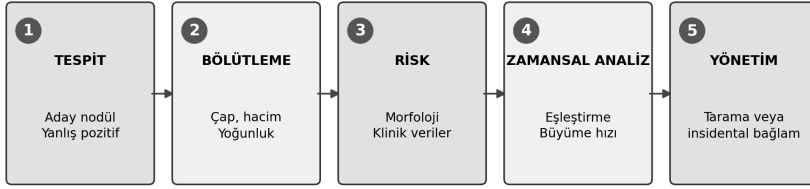
özelliklerini yüksek boyutlu sayısal değişkenlere dönüştürür. Bu yöntemler risk sınıflamasına katkı sağlayabilse de geliştirildikleri popülasyonun prevalansına ve referans standardına bağımlıdır (Ardila vd., 2019; Wu vd., 2022; Liu vd., 2022).

Radyomik özellikler segmentasyon yönteminden, kesit kalınlığından, rekonstrüksiyon algoritmasından ve üreticiye özgü görüntü işleme adımlarından etkilenebilir. İç doğrulamada yüksek ayırt etme performansı gösteren bir modelin dış merkezde kalibrasyonu bozulabilir. Bu nedenle radyomik ve radyogenomik yaklaşımlar günümüzde özellikle araştırma ve risk zenginleştirme araçları olarak görülmeli; tek başına biyopsi veya cerrahi kararının gerekçesi haline getirilmemelidir (Magnin vd., 2025; Boubnovski Martell vd., 2025).

Zamansal İzlem ve Büyüme Öngörüsü

Seri BT incelemelerinin otomatik eşleştirilmesi, aynı nodülün yeniden bulunması ve hacim değişiminin hesaplanması radyoloğun iş yükünü azaltabilir. Rutin klinik popülasyonda değerlendirilen bazı ticari sistemlerin tespit, sınıflandırma, hacim ve büyüme hızı ölçümünde yardımcı olabileceği bildirilmiştir; ancak yanlış eşleştirme ve bölütleme hataları radyolog denetimini gerekli kılar (Murchison vd., 2022).

Başlangıç BT'sinden nodülün gelecekteki büyümesini veya takip görüntüsünü öngörmeye çalışan modeller geliştirilmiş olmakla birlikte henüz daha sınırlıdır. Çok görevli öğrenme yaklaşımıyla yayımlanan güncel bir çalışma umut verici sonuçlar bildirmiştir; ancak tek merkezli 246 hastaya, iç çapraz doğrulamaya ve ağırlıklı olarak görüntü benzerliği ile bölütleme ölçütlerine dayanmaktadır. Dış doğrulama, prospektif klinik değerlendirme ve hasta sonuçlarıyla ilişki gösterilmeden bu tür modeller klinik izlem aralığını değiştiren araçlar olarak kullanılmamalıdır (Zhao vd., 2026). Güncel derlemeler de veri heterojenliği, standart büyüme tanımı ve klinik geçerlilik sorunlarının sürdüğünü vurgulamaktadır (Huang vd., 2026).



Her aşama radyolog doğrulaması, önceki incelemeler ve uygun yönetim kılavuzu ile birlikte değerlendirilir.

Şekil 2. Pulmoner nodül yönetiminde yapay zekanın destekleyebileceği ardışık görevler.

DIFFÜZ PARANKİMAL AKCİĞER HASTALIKLARINDA YAPAY ZEKA

Patern Tanımlama ve Nicel Değerlendirme

Diffüz parankimal akciğer hastalıkları, interstisyel akciğer hastalıklarını da içeren geniş bir hastalık grubudur. Yüksek çözünürlüklü BT'de buzlu cam, retikülasyon, traksiyon bronşektazisi, bal peteği ve amfizem gibi bulguların dağılımı tanısal değerlendirmede önem taşır. YZ sistemleri lokal parankimal paternleri sınıflandırabilir, etkilenen akciğer hacmini nicel olarak hesaplayabilir veya benzer görüntülere sahip olguları veri tabanından geri getirebilir (Kim vd., 2018; Christe vd., 2019; Choe vd., 2022).

Lokal patern sınıflandırması ile hastalık düzeyinde tanı aynı görev değildir. Olağan interstisyel pnömoni, nonspesifik interstisyel pnömoni veya fibrotik hipersensitivite pnömonisi gibi tanılar; paternin yanı sıra dağılım, klinik öykü, maruziyet, seroloji ve gerektiğinde patoloji verilerinin birlikte değerlendirilmesini gerektirir. Görüntü temelli bir modelin yüksek sınıflandırma başarısı, multidisipliner tanı sürecinin yerini aldığı biçiminde yorumlanmamalıdır (Koo vd., 2022; Exarchos vd., 2023).

Nicel BT analizi, hastalık yükünü sürekli bir değişken olarak sunarak görsel yarı nicel skorlamayı tamamlayabilir. Bununla birlikte akciğer hacmi, inspirasyon düzeyi, hareket, rekonstrüksiyon yöntemi ve eşlik eden enfeksiyon veya ödem gibi durumlar ölçümü etkileyebilir. Seri değerlendirmelerde teknik standardizasyon ve algoritmanın aynı sürümünün kullanılması önemlidir.

Progresyon ve Prognoz

İdiyopatik pulmoner fibrozis ve diğer fibrotik interstisyel akciğer hastalıklarında erken nicel değişikliklerin solunum fonksiyonları ve klinik seyirle ilişkili olabileceği bildirilmiştir (Kim vd., 2020; Aoki vd., 2022). Bu yaklaşım, görsel olarak sınırlı değişiklik gösteren olgularda hastalık yükünün daha tekrarlanabilir izlenmesine katkı sağlayabilir. Ancak korelasyon gösterilmesi, model kullanımının tedavi kararını veya sağkalımı iyileştirdiğini kanıtlamaz.

Prognostik modellerin klinik uygulamaya geçebilmesi için tedavi, hastalık şiddeti ve rekabet eden risklerin dikkate alındığı çok merkezli dış doğrulamalara ihtiyaç vardır. Model çıktısı mevcut klinik göstergelere ek değer sağlamalı, iyi kalibre edilmeli ve karar eşiğinin hasta yönetiminde hangi değişikliği oluşturacağı önceden tanımlanmalıdır.

TRİYAJ VE İŞ LİSTESİ ÖNCELİKLENDİRMESİ

YZ destekli triyaj sistemleri, görüntü elde edildikten sonra belirli bir kritik bulgu açısından incelemeyi tarar ve şüpheli olguyu iş listesinde öne taşır. Akciğer grafisinde pnömotoraks veya belirgin plevral efüzyon, BT pulmoner anjiyografide pulmoner emboli bu yaklaşımın başlıca örnekleridir. Bu sistemlerin amacı nihai tanı koymak değil, radyolog değerlendirmesine kadar geçen süreyi kısaltmaktır (Hwang vd., 2025). Akciğer grafisi akut aortik sendromu dışlayamaz; bu nedenle grafideki yardımcı bulguların işaretlenmesi, hastalığın tanısı olarak sunulmamalıdır.

Akciğer grafilerinin otomatik önceliklendirilmesine ilişkin çalışmalarda kritik incelemelerin raporlanma süresinin simülasyon ortamında kısalabildiği gösterilmiştir (Annarumma vd., 2019; Baltruschat vd., 2021). Bununla birlikte simülasyon sonucu, gerçek zamanlı klinik uygulama veya hasta sonucuyla eşdeğer değildir. Modelin yanlış negatif olarak düşük önceliğe yerleştirdiği incelemeler gecikebilir; normal veya düşük öncelikli gruptaki bekleme süresi uzayabilir.

Güvenli bir triyaj iş akışında hedef bulgu, kabul edilebilir en uzun bekleme süresi, uyarının kime ve nasıl iletileceği, alarm tekrarı, sistem kesintisinde geri dönüş planı ve yanlış alarm yükü önceden tanımlanmalıdır. Performans yalnızca duyarlılık ve özgüllükle değil; görüntüden rapora ve görüntüden klinik müdahaleye kadar geçen süre, kaçırılan kritik olgu sayısı ve uyarı yorgunluğu ile birlikte izlenmelidir.

DÜŞÜK DOZ BT VE DERİN ÖĞRENME REKONSTRÜKSİYONU

Toraks BT'de radyasyon dozu ile görüntü kalitesi arasındaki denge özellikle tarama ve uzun süreli takip gerektiren hastalarda önemlidir. Filtrelenmiş geri projeksiyon düşük dozda gürültülü görüntüler üretebilir; iteratif rekonstrüksiyon gürültüyü azaltırken görüntü dokusunu değiştirebilir. Derin öğrenme tabanlı rekonstrüksiyon (deep learning image reconstruction; DLIR) yöntemleri, gürültüyü baskımlarken uzaysal çözünürlüğü ve daha doğal görüntü dokusunu korumayı amaçlar (Koetzier vd., 2023).

DLIR çalışmalarında görüntü gürültüsü, sinyal-gürültü oranı, öznel kalite, nodül görünürlüğü ve interstisyel patern değerlendirmesi açısından olumlu sonuçlar bildirilmiştir. Bununla birlikte görüntü kalitesindeki iyileşme, tüm klinik görevlerde tanısal üstünlük anlamına gelmez. Güncel sistematik derlemede çalışmaların protokol, doz, üretici ve sonlanım ölçütleri açısından heterojen olduğu; hasta sonuçlarını gösterecek geniş ölçekli çalışmalara gereksinim bulunduğı belirtilmektedir (Obhuli vd., 2026).

DLIR etkisi üreticiye, model gücüne, hasta yapısına, kesit kalınlığına ve rekonstrüksiyon çekirdeğine bağlıdır. Aşırı gürültü baskılama küçük buzlu cam nodüllerini veya ince retikülasyonu yumuşatabilir; doku özelliklerinin değişmesi radyomik analizlerin tekrarlanabilirliğini etkileyebilir. Bu nedenle doz azaltımı kurumun standart protokolü, hedef klinik görev ve yerel fantom/olgu değerlendirmesi temelinde yapılmalı; yalnızca üreticinin önerdiği sayısal oranlara dayandırılmamalıdır.

KLİNİK KANIT, ENTEGRASYON VE SINIRLILIKLAR

Kanıt Düzeyinin Yorumlanması

Toraks radyolojisinde YZ kanıtı homojen değildir. Teknik geliştirme çalışmaları genellikle seçilmiş veri kümelerinde iç test performansını bildirir. Dış doğrulama farklı merkez veya popülasyondaki genellenebilirliği, okuyucu çalışmaları radyologla etkileşimi, prospektif iş akışı çalışmaları gerçek kullanım etkisini ve hasta sonucu çalışmaları klinik yararı değerlendirir. Bir alt basamaktaki başarı, üst basamaktaki yararın kanıtı olarak kabul edilmemelidir.

Mevcut kanıtlar akciğer grafisinde dar görevli okuyucu desteğı, tarama BT'sinde nodül tespiti/ölçümü ve bazı triyaj uygulamalarında birikmektedir. İnterstisyel hastalıkların nicel değerlendirilmesi ve DLIR gelişmekte olan alanlardır. Başlangıç görüntüsünden nodül büyümesi, moleküler özellik veya

hasta prognozu öngören yüksek boyutlu modeller ise çoğunlukla araştırma aşamasındadır (Hwang vd., 2025; Huang vd., 2026).

Tablo 1. Toraks radyolojisinde başlıca yapay zeka uygulamalarının görelî kanıt olgunluğu

Uygulama alanı	Temel görev	Kanıt durumu	Başlıca sınırlılık
Akciğer grafisi	Anormallik tespiti ve okuyucu desteği	Orta	Cihaz, projeksiyon, prevalans ve eşik değerine bağlı performans; yanlış uyarılar
Tarama BT	Nodül tespiti, ölçüm ve karşılaştırma	Orta	Benign nodül yükü; kanser saptama ve hasta sonucu katkısı sınırlı
İnsidental nodül	Risk sınıflaması ve yönetim desteği	Gelişmekte	Tarama modellerinin farklı popülasyona aktarılması; kalibrasyon
Nodül büyüme öngörüsü	Başlangıç BT'sinden gelecekteki değişimin tahmini	Deneysel	Tek merkezli veri, iç doğrulama ve klinik sonuç eksikliği
Diffüz parankimal hastalık	Patern sınıflaması ve nicel hastalık yükü	Gelişmekte	Teknik standardizasyon; hastalık tanısı için klinik bağlam gereksinimi
Triyaj	İş listesi önceliklendirmesi	Orta	Yanlış negatif olgunun gecikmesi, alarm yorgunluğu, iş akışı bağımlılığı
DLIR	Gürültü azaltma ve doz optimizasyonu	Gelişmekte	Üretici/protokol bağımlılığı; doku değişimi ve hasta sonucu kanıtı

Not: Kanıt durumu, uygulamanın genel literatür olgunluğunu nitel olarak özetlemektedir; ürün ve kurum düzeyindeki performans ayrıca doğrulanmalıdır. DLIR: derin öğrenme tabanlı rekonstrüksiyon; BT: bilgisayarlı tomografi.

Veri Kayması, Kalibrasyon ve Yerel Doğrulama

YZ performansı eğitim verisinin temsil gücüne bağlıdır. Akciğer grafisinde posteroanterior ve taşınabilir anteroposterior çekimler, yoğun bakım cihazları ve farklı hasta profilleri; BT'de kesit kalınlığı, doz, kontrast fazı, rekonstrüksiyon algoritması ve üretici farklılıkları alan kaymasına yol açabilir. Belirli coğrafya ve demografik grupların aşırı temsil edilmesi de alt gruplar arasında performans eşitsizliği oluşturabilir (Tripathi vd., 2023; Hwang vd., 2025).

Bir modelin alan altındaki eğrisi yüksek olsa bile seçilen çalışma eşğinde pozitif ve negatif öngörü değerleri hastalık prevalansı ile değişir. Klinik uygulama öncesinde yerel ve ardışık olgulardan oluşan bir veri kümesinde duyarlılık, özgüllük, yanlış uyarı/inceleme, kalibrasyon ve alt grup performansı değerlendirilmelidir. İlk aşamada sonuçların klinisyene gösterilmediği sessiz doğrulama, iş akışı etkisini ve teknik hataları hasta bakımını etkilemeden belirlemeye yardımcı olabilir.

Referans Standardı ve İnsan-YZ Etkileşimi

Eğitim etiketleri radyoloji raporlarından, tek okuyucudan, uzman uzlaşısından veya patoloji ve takip sonuçlarından üretilebilir. Bu referanslar aynı güvenilirlik düzeyine sahip değildir. Toraks görüntülemesinde subjektif paternler ve eksik rapor etiketleri modelin aynı hataları öğrenmesine neden olabilir. Model performansı değerlendirilirken etiketleme yöntemi, okuyucu sayısı, uzlaşısı süreci ve belirsiz olguların nasıl ele alındığı açıkça bildirilmelidir (Calandriello ve Walsh, 2021).

Radyoloğun YZ çıktısına aşırı güvenmesi otomasyon yanlılığına, çıktıyı sistematik biçimde göz ardı etmesi ise beklenen yararın kaybolmasına yol açabilir. Arayüz, modelin olasılık değerini ve işaretlediği bölgeyi anlaşılır biçimde göstermeli; ancak ısı haritası veya dikkat haritası modelin kararının nedensel açıklaması ya da doğruluk kanıtı olarak kabul edilmemelidir. Açıklanabilirlik yöntemleri hata analizi, kalibrasyon, dış doğrulama ve kullanıcı eğitiminin yerine geçmez (Choo ve Liu, 2018; Liu vd., 2022).

Klinik Entegrasyon ve Sürekli İzlem

Klinik uygulama yalnızca algoritmanın PACS'a bağlanması değildir. Hangi incelemelerin modele gönderileceği, sonuçların ne kadar sürede döneceği, hatalı veya eksik analizlerin nasıl gösterileceği, kullanıcıların kim olduğu ve sonucun rapora nasıl aktarılacağı tanımlanmalıdır. Veri güvenliği, erişim yetkileri, kayıt izi, bakım sorumluluğu ve sistem kesintisinde standart iş listesine dönüş planı entegrasyonun parçasıdır (Hwang ve Park, 2020; Hwang vd., 2025).

Cihaz, protokol, popülasyon veya yazılım sürümü değiştiğinde performans zaman içinde kayabilir. Bu nedenle uygulama sonrasında yanlış pozitif/negatif oranları, rapor değişikliği, sonuçlanma süresi, kullanıcı kabulü ve alt grup performansı düzenli aralıklarla izlenmelidir. Büyük yazılım güncellemeleri yeni bir doğrulama gerektirebilir. Nihai klinik sorumluluk ve olağandışı sonuçlarda izlenecek süreç kurum politikalarında açıkça belirtilmelidir.

Eşitlik ve Çok Modlu Modeller

Ticari ürünlere erişim, bilgi teknolojisi altyapısı ve yerel veri oluşturma kapasitesi kurumlar arasında farklılık gösterebilir. Eşitsiz temsil edilen popülasyonlarda düşük performans, mevcut sağlık eşitsizliklerini artırabilir. Görüntüleri klinik öykü, laboratuvar, solunum fonksiyon testi ve önceki raporlarla birleştiren çok modlu modeller daha bütüncül sonuçlar sunma

potansiyeline sahiptir; ancak veri eksikliği, mahremiyet, önyargı ve doğrulama gereksinimleri daha da karmaşık hale gelir. Bu modeller için de dar ve açık kullanım amacı, dış doğrulama ve insan denetimi temel koşullardır.

SONUÇ VE ÖNERİLER

Toraks radyolojisinde YZ için en olgun kullanım alanları, akciğer grafisinde belirli anormalliklerin ikinci okuyucu olarak tespiti, tarama BT'sinde nodül tespiti ve ölçümü ile bazı triyaj görevleridir. Bu uygulamalar radyolog performansını ve iş akışını destekleyebilir; ancak yararın büyüklüğü ürün, kurum, hasta grubu ve kullanım biçimine bağlıdır.

Diffüz parankimal hastalıklarda nicel değerlendirme ve DL rekonstrüksiyonu geliştirmekte olan alanlardır. Radyomik risk modelleri, başlangıç görüntüsünden nodül büyüme öngörüsü ve çok modlu prognostik sistemler ise umut verici olmakla birlikte çoğunlukla araştırma aşamasındadır. Teknik doğruluk; klinik yarar, maliyet etkinliği veya hasta sonucu ile eş anlamlı değildir.

Güvenli uygulama için modelin kullanım amacı tanımlanmalı, yerel veride doğrulanmalı, iş akışına kontrollü biçimde entegre edilmeli ve sonrasında izlenmelidir. YZ radyoloğun uzmanlığını tamamlamalı; nihai yorum ve hasta yönetimi insan denetiminde kalmalıdır.

KAYNAKÇA

- Aberle, D. R., Adams, A. M., Berg, C. D., Black, W. C., Clapp, J. D., Fagerstrom, R. M., Gareen, I. F., Gatsonis, C., Marcus, P. M. ve Sicks, J. D. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. *The New England Journal of Medicine*, 365(5), 395-409. <https://doi.org/10.1056/NEJMoa1102873>
- Annarumma, M., Withey, S. J., Bakewell, R. J., Pesce, E., Goh, V. ve Montana, G. (2019). Automated triaging of adult chest radiographs with deep artificial neural networks. *Radiology*, 291(1), 196-202. <https://doi.org/10.1148/radiol.2018180921>
- Aoki, R., Iwasawa, T., Saka, T., Yamashiro, T., Utsunomiya, D., Misumi, T., Baba, T. ve Ogura, T. (2022). Effects of automatic deep-learning-based lung analysis on quantification of interstitial lung disease: Correlation with pulmonary function test results and prognosis. *Diagnostics*, 12(12), 3038. <https://doi.org/10.3390/diagnostics12123038>
- Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G., Naidich, D. P. ve Shetty, S. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine*, 25(6), 954-961. <https://doi.org/10.1038/s41591-019-0447-x>
- Armato, S. G. III, McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., Zhao, B., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., MacMahon, H., van Beek, E. J. R., Yankelevitz, D., Biancardi, A. M., Bland, P. H., Brown, M. S., Engelmann, R. M., Laderach, G. E., ... Croft, B. Y. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*, 38(2), 915-931. <https://doi.org/10.1118/1.3528204>
- Baltruschat, I., Steinmeister, L., Nickisch, H., Saalbach, A., Grass, M., Adam, G., Knopp, T. ve Ittrich, H. (2021). Smart chest X-ray worklist prioritization using artificial intelligence: A clinical workflow simulation. *European Radiology*, 31(6), 3837-3845. <https://doi.org/10.1007/s00330-020-07480-7>
- Bennani, S., Regnard, N. E., Ventre, J., Lassalle, L., Nguyen, T., Ducarouge, A., Dargent, L., Guillo, E., Gouhier, E., Zaimi, S. H., Canniff, E., Malandrin, C., Khafagy, P., Koulakian, H., Revel, M. P. ve Chassagnon, G. (2023). Using AI to improve radiologist performance in detection of abnormalities on chest radiographs. *Radiology*, 309(3), e230860. <https://doi.org/10.1148/radiol.230860>
- Boubnovski Martell, M., Linton-Reid, K., Chen, M. ve Aboagye, E. O. (2025). Radiomics for lung cancer diagnosis, management, and future prospects. *Clinical Radiology*, 86, 106926. <https://doi.org/10.1016/j.crad.2025.106926>
- Calandriello, L. ve Walsh, S. L. F. (2021). Artificial intelligence for thoracic radiology: From research tool to clinical practice. *European Respiratory Journal*, 57(5), 2100625. <https://doi.org/10.1183/13993003.00625-2021>
- Chen, H., Liu, J., Lu, L., Wang, T., Xu, X., Chu, A., Peng, W., Gong, J., Tang, W. ve Gu, Y. (2022). Volumetric segmentation of ground-glass nodule based on 3D attentional cascaded residual U-Net and conditional random field. *Medical Physics*, 49(2), 1097-1107. <https://doi.org/10.1002/mp.15423>
- Choe, J., Hwang, H. J., Seo, J. B., Lee, S. M., Yun, J., Kim, M. J., Jeong, J., Lee, Y., Jin, K., Park, R., Kim, J., Jeon, H., Kim, N., Yi, J., Yu, D. ve Kim, B. (2022). Content-based image retrieval by using deep learning for interstitial lung disease diagnosis with chest CT. *Radiology*, 302(1), 187-197. <https://doi.org/10.1148/radiol.2021204164>
- Choo, J. ve Liu, S. (2018). Visual analytics for explainable deep learning. *IEEE Computer Graphics and Applications*, 38(4), 84-92. <https://doi.org/10.1109/MCG.2018.042731661>
- Christe, A., Peters, A. A., Drakopoulos, D., Heverhagen, J. T., Geiser, T., Stathopoulou, T., Christodoulidis, S., Anthimopoulos, M., Mougiakakou, S. G. ve Ebner, L. (2019). Computer-aided diagnosis of pulmonary fibrosis using deep learning and CT images. *Investigative Radiology*, 54(10), 627-632. <https://doi.org/10.1097/RLI.0000000000000574>
- Christensen, J., Prosper, A. E., Wu, C. C., Chung, J., Lee, E., Elicker, B., Hunsaker, A. R., Petranovic, M., Sandler, K. L., Stiles, B., Mazzone, P., Yankelevitz, D., Aberle, D., Chiles, C. ve Kazerooni, E. (2024). ACR Lung-RADS v2022: Assessment categories and management recommendations. *Chest*, 165(3), 738-753. <https://doi.org/10.1016/j.chest.2023.10.028>

- Exarchos, K. P., Gkrepi, G., Kostikas, K. ve Gogali, A. (2023). Recent advances of artificial intelligence applications in interstitial lung diseases. *Diagnostics*, 13(13), 2303. <https://doi.org/10.3390/diagnostics13132303>
- Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. ve Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, 18(8), 500-510. <https://doi.org/10.1038/s41568-018-0016-5>
- Hua, K. L., Hsu, C. H., Hidayati, S. C., Cheng, W. H. ve Chen, Y. J. (2015). Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and Therapy*, 8, 2015-2022. <https://doi.org/10.2147/OTT.S80733>
- Huang, W., Ye, M., Wang, B., Li, Q., Luo, J., Wen, M., Chen, Y., Jing, W. ve Fan, L. (2026). Narrative review: The research advances of artificial intelligence in the prediction of pulmonary nodule growth. *Journal of Thoracic Disease*, 18(2), 165. <https://doi.org/10.21037/jtd-2025-1580>
- Hwang, E. J., Goo, J. M. ve Park, C. M. (2025). AI applications for thoracic imaging: Considerations for best practice. *Radiology*, 314(2), e240650. <https://doi.org/10.1148/radiol.240650>
- Hwang, E. J. ve Park, C. M. (2020). Clinical implementation of deep learning in thoracic radiology: Potential applications and challenges. *Korean Journal of Radiology*, 21(5), 511-525. <https://doi.org/10.3348/kjr.2019.0821>
- Hwang, E. J., Park, S., Jin, K. N., Kim, J. I., Choi, S. Y., Lee, J. H., Goo, J. M., Aum, J., Yim, J. J., Cohen, J. G., Ferretti, G. R. ve Park, C. M. (2019). Development and validation of a deep learning-based automated detection algorithm for major thoracic diseases on chest radiographs. *JAMA Network Open*, 2(3), e191095. <https://doi.org/10.1001/jamanetworkopen.2019.1095>
- Kim, G. B., Jung, K. H., Lee, Y., Kim, H. J., Kim, N., Jun, S., Seo, J. B. ve Lynch, D. A. (2018). Comparison of shallow and deep learning methods on classifying the regional pattern of diffuse lung disease. *Journal of Digital Imaging*, 31(4), 415-424. <https://doi.org/10.1007/s10278-017-0028-9>
- Kim, G. H. J., Weigt, S. S., Belperio, J. A., Brown, M. S., Shi, Y., Lai, J. H. ve Goldin, J. G. (2020). Prediction of idiopathic pulmonary fibrosis progression using early quantitative changes on CT imaging for a short term of clinical 18-24-month follow-ups. *European Radiology*, 30(2), 726-734. <https://doi.org/10.1007/s00330-019-06402-6>
- Koetzier, L. R., Mastrodicasa, D., Szczykutowicz, T. P., van der Werf, N. R., Wang, A. S., Sandfort, V., van der Molen, A. J., Fleischmann, D. ve Willeminck, M. J. (2023). Deep learning image reconstruction for CT: Technical principles and clinical prospects. *Radiology*, 306(3), e221257. <https://doi.org/10.1148/radiol.221257>
- Koo, C. W., Williams, J. M., Liu, G., Panda, A., Patel, P. P., Frota Lima, L. M. M., Karwoski, R. A., Moua, T., Larson, N. B. ve Bratt, A. (2022). Quantitative CT and machine learning classification of fibrotic interstitial lung diseases. *European Radiology*, 32(12), 8152-8161. <https://doi.org/10.1007/s00330-022-08875-4>
- Lee, S. M., Seo, J. B., Yun, J., Cho, Y. H., Vogel-Claussen, J., Schiebler, M. L., Gefter, W. B., van Beek, E. J. R., Goo, J. M., Lee, K. S., Hatabu, H., Gee, J. ve Kim, N. (2019). Deep learning applications in chest radiography and computed tomography: Current state of the art. *Journal of Thoracic Imaging*, 34(2), 75-85. <https://doi.org/10.1097/RTI.0000000000000387>
- Liu, J. A., Yang, I. Y. ve Tsai, E. B. (2022). Artificial intelligence for lung nodules: From the AJR special series on AI applications. *AJR American Journal of Roentgenology*, 219(5), 703-712. <https://doi.org/10.2214/AJR.22.27487>
- Ma, J., Song, Y., Tian, X., Hua, Y., Zhang, R. ve Wu, J. (2020). Survey on deep learning for pulmonary medical imaging. *Frontiers of Medicine*, 14(4), 450-469. <https://doi.org/10.1007/s11684-019-0726-4>
- MacMahon, H., Naidich, D. P., Goo, J. M., Lee, K. S., Leung, A. N. C., Mayo, J. R., Mehta, A. C., Ohno, Y., Powell, C. A., Prokop, M., Rubin, G. D., Schaefer-Prokop, C. M., Travis, W. D., Van Schil, P. E. ve Bankier, A. A. (2017). Guidelines for management of incidental pulmonary nodules detected on CT images: From the Fleischner Society 2017. *Radiology*, 284(1), 228-243. <https://doi.org/10.1148/radiol.2017161659>

- Magnin, C. Y., Lauer, D., Ammeter, M. ve Gote-Schniering, J. (2025). From images to clinical insights: An educational review on radiomics in lung diseases. *Breathe*, 21(1), 230225. <https://doi.org/10.1183/20734735.0225-2023>
- Min, Y., Hu, L., Wei, L. ve Nie, S. (2022). Computer-aided detection of pulmonary nodules based on convolutional neural networks: A review. *Physics in Medicine & Biology*, 67(6), 06TR01. <https://doi.org/10.1088/1361-6560/ac568e>
- Murchison, J. T., Ritchie, G., Senyszak, D., Nijwening, J. H., van Veenendaal, G., Wakkie, J. ve van Beek, E. J. R. (2022). Validation of a deep learning computer aided system for CT based lung nodule detection, classification, and growth rate estimation in a routine clinical population. *PLOS ONE*, 17(5), e0266799. <https://doi.org/10.1371/journal.pone.0266799>
- Nam, J. G., Kim, M., Park, J., Hwang, E. J., Lee, J. H., Hong, J. H., Goo, J. M. ve Park, C. M. (2021). Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *European Respiratory Journal*, 57(5), 2003061. <https://doi.org/10.1183/13993003.03061-2020>
- Obhuli, C. M., Pendem, S., Abhijith, S., Kadavigere, R., Priyanka, Priya, P. S. ve Chacko, C. (2026). Clinical value of deep learning image reconstruction in chest computed tomography imaging: A systematic review. *Clinical Radiology*, 92, 107184. <https://doi.org/10.1016/j.crad.2025.107184>
- Setio, A. A. A., Traverso, A., de Bel, T., Berens, M. S. N., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M. E., Geurts, B., van der Gugten, R., Heng, P. A., Jansen, B., de Kaste, M. M. J., Kotov, V., Lin, J. Y., Manders, J. T. M. C., Sonora-Mengana, A., Garcia-Naranjo, J. C., ... Jacobs, C. (2017). Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Medical Image Analysis*, 42, 1-13. <https://doi.org/10.1016/j.media.2017.06.015>
- Shin, H. J., Han, K., Ryu, L. ve Kim, E. K. (2023). The impact of artificial intelligence on the reading times of radiologists for chest radiographs. *npj Digital Medicine*, 6(1), 82. <https://doi.org/10.1038/s41746-023-00829-4>
- Tripathi, S., Gabriel, K., Dheer, S., Parajuli, A., Augustin, A. I., Elahi, A., Awan, O. ve Dako, F. (2023). Understanding biases and disparities in radiology AI datasets: A review. *Journal of the American College of Radiology*, 20(9), 836-841. <https://doi.org/10.1016/j.jacr.2023.06.015>
- Wu, Z., Wang, F., Cao, W., Qin, C., Dong, X., Yang, Z., Zheng, Y., Luo, Z., Zhao, L., Yu, Y., Xu, Y., Li, J., Tang, W., Shen, S., Wu, N., Tan, F., Li, N. ve He, J. (2022). Lung cancer risk prediction models based on pulmonary nodules: A systematic review. *Thoracic Cancer*, 13(5), 664-677. <https://doi.org/10.1111/1759-7714.14333>
- Zhao, W., Chen, Y., Xie, Y., Nie, S., Han, B., Jiang, Y. ve Li, X. (2026). Multi-task learning for predicting pulmonary nodule growth and follow-up volume. *Frontiers in Oncology*, 16, 1677521. <https://doi.org/10.3389/fonc.2026.1677521>

MEME RADYOLOJİSİNDE YAPAY ZEKA: TARAMA VE TANIDA KARAR DESTEK SİSTEMLERİ

Derya Güçlü*

GİRİŞ

Meme kanseri, kadınlarda en sık tanı alan malignitelerden biridir. Prognozun iyileştirilmesinde erken tanı belirleyici olduğundan, nüfus temelli tarama programlarının temel yöntemi mamografidir. Dijital meme tomosentezi (DBT), ultrasonografi (US) ve manyetik rezonans görüntüleme (MRG) ise tarama ya da tanısal değerlendirmede klinik risk, meme yoğunluğu ve bulguya göre tamamlayıcı roller üstlenir. Meme görüntülerinin yorumlanması; örtüşen parankim, ince yapısal distorsiyonlar, mikrokalsifikasyon kümeleri ve gözlemciler arası değişkenlik nedeniyle yüksek dikkat gerektirir. Tarama hacmindeki artış da bu güçlüğe iş yükü boyutunu eklemektedir (Abu Abeelh ve Abuabeileh, 2025; Díaz vd., 2024).

Bilgisayar destekli saptama sistemleri meme görüntülemeye yeni değildir. İlk kuşak CAD yazılımları elle tanımlanan görüntü özelliklerine dayanıyor ve şüpheli bölgeleri radyoloğa işaretliyordu. Bununla birlikte geniş ölçekli klinik değerlendirmelerde geleneksel CAD kullanımının tanısal doğruluğu belirgin biçimde artırmadığı, yanlış pozitif işaretlemeleri ve geri çağırmaı artırabildiği gösterildi (Lehman vd., 2015). Güncel yapay zeka (YZ) sistemleri ise özellikleri veriden öğrenen derin sinir ağlarını kullanır; bu nedenle önceki CAD kuşağından teknik olarak farklıdır. Ancak daha yüksek teknik performans, tek başına klinik yarar anlamına gelmez. Klinik etki; modelin hangi iş akışına yerleştirildiğine, karar eşiğine, tarama popülasyonuna ve radyolog-algoritma etkileşimine bağlıdır.

Meme görüntülemeye en yaygın modeller, konvolüsyonel sinir ağları ile görüntü düzeyinde kanser olasılığı üretir ve çoğu zaman şüpheli alanları ısı haritası veya işaretleme ile gösterir. Segmentasyon için U-Net benzeri mimariler, sınıflandırma için ResNet ve DenseNet türevleri kullanılabilir.

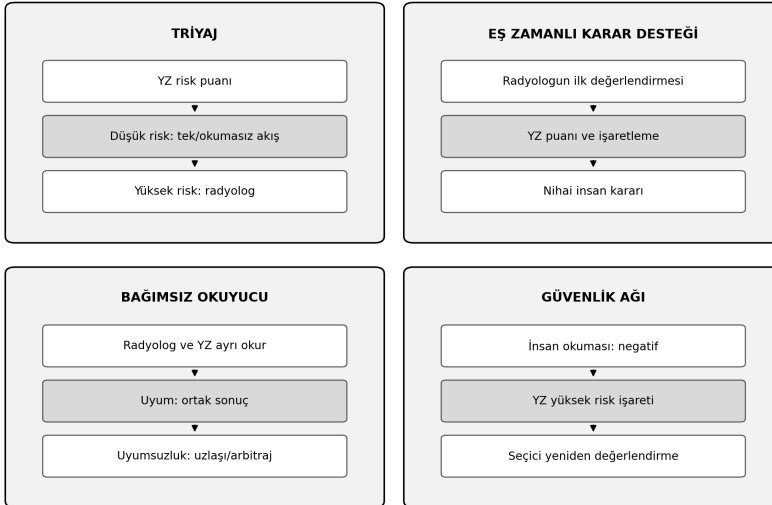
* Doç. Dr., Düzce Üniversitesi, Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, deryasr@hotmail.com, ORCID: <https://orcid.org/0000-0001-5332-2909>

DBT’de üç boyutlu görüntü hacminin işlenmesi, US’de cihaz ve operatör değişkenliği, MRG’de ise çok fazlı dinamik kontrastlı verinin uzamsal-zamansal analizi ayrı teknik gereksinimler doğurur. Bu nedenle “meme görüntüleme YZ’si” tek bir ürün veya tek bir klinik görev değildir; saptama, karakterizasyon, yoğunluk ölçümü, risk tahmini ve iş akışı yönetimi birbirinden ayrılmalıdır.

MAMOGRAFİDE YAPAY ZEKA

Klinik Kullanım Modelleri

Mamografi taramasında YZ dört temel biçimde kullanılabilir. Triyaj yaklaşımında algoritma incelemeleri risk puanına göre ayırır; düşük riskli incelemeler tek okumaya yönlendirilebilir veya önceden tanımlanmış koşullarda insan okumasından çıkarılabilir. Eş zamanlı karar desteğinde radyolog önce görüntüyü değerlendirir, ardından YZ puanı ve işaretlemelerini kullanarak nihai kararını verir. Bağımsız okuyucu modelinde YZ, çift okuma sistemindeki insan okuyuculardan birinin yerini alır ve uyumsuzluk durumunda uzlaşa ya da arbitraj yapılır. Güvenlik ağı modelinde ise insanlar tarafından negatif değerlendirilen incelemeler arasından YZ’nin yüksek riskli buldukları seçici olarak yeniden incelenir. Bu modellerin iş yükü, kanser saptama ve geri çağırma üzerindeki etkileri aynı değildir; sonuçlar kullanım modeli belirtilmeden karşılaştırılmamalıdır.



Şekil 1. Tarama mamografisinde başlıca yapay zeka kullanım modelleri. Modeller birbirinin eş anlamlısı değildir; performans ve sorumluluk değerlendirmesi seçilen iş akışına göre yapılmalıdır.

Retrospektif Performanstan Prospektif Kanıt

Geniş retrospektif çalışmalar, derin öğrenme modellerinin belirli veri setlerinde deneyimli radyologlara yakın performans gösterebildiğini ortaya koymuştur. Rodriguez-Ruiz ve arkadaşlarının 101 radyologla yaptığı karşılaştırmada bağımsız YZ sisteminin doğruluğu ortalama radyolog performansına yakın bulunmuştur. McKinney ve arkadaşlarının Birleşik Krallık ve ABD verileriyle yürüttüğü uluslararası değerlendirme de yanlış pozitif ve yanlış negatif oranlarında azalma potansiyeli bildirmiştir (Rodriguez-Ruiz vd., 2019; McKinney vd., 2020). Bu sonuçlar teknolojinin klinik araştırmaya taşınmasını sağlamış olsa da retrospektif tasarım, seçilmiş veri kümeleri ve önceden belirlenmiş eşikler nedeniyle doğrudan gerçek yaşam yararı olarak yorumlanamaz.

Prospektif çalışmalar YZ'nin tarama iş akışındaki rolünü daha somutlaştırmıştır. ScreenTrustCAD çalışmasında 55.581 kadının taramasında bir radyolog ile bağımsız YZ okuyucusunun oluşturduğu çift, standart iki radyologlu okumaya kanser saptama bakımından aşağı kalmamış; iki yaklaşım sırasıyla 261 ve 250 kanser saptamıştır (Dembrower vd., 2023). Buna karşılık Ng ve arkadaşlarının uygulamasında YZ bir insan okuyucunun yerine geçmemiş, standart çift okumada negatif kabul edilen olguları ek okuyucu olarak işaretlemiştir. Bu güvenlik ağı yaklaşımı her 1.000 incelemede yaklaşık 0,7–1,6 ek kanser saptarken ek geri çağırma oranı %0,16–0,30 düzeyinde kalmıştır (Ng vd., 2023). Dolayısıyla her iki çalışma “YZ destekli okuma” başlığı altında anılsa da biri okuyucu ikamesini, diğeri ek güvenlik katmanını değerlendirmiştir.

MASAI randomize çalışmasının ilk güvenlik analizi YZ destekli stratejinin okuma iş yükünü yaklaşık %44 azalttığını ve kanser saptamayı artırabildiğini göstermiştir (Lång vd., 2023; Hernström vd., 2025). 2026 tarihli sonuçlarda 105.934 kadın değerlendirilmiş; interval kanser oranı YZ destekli grupta 1,55/1.000, standart çift okumada 1,76/1.000 bulunmuş ve aşağı kalmama ölçütü karşılanmıştır. Duyarlılık %80,5'e karşı %73,8 ile YZ grubunda daha yüksek, özgüllük ise her iki grupta %98,5 olmuştur. Sonuçların yaş ve meme yoğunluğu alt gruplarında genel olarak tutarlı olması önemlidir; buna karşın çalışma tek ülke, belirli cihazlar ve tek bir ticari sistem bağlamında yürütülmüştür (Gommers vd., 2026).

Gerçek yaşam verileri de giderek güçlenmektedir. Almanya'daki gözlemsel PRAIM uygulamasında 463.094 kadın ve 119 radyolog değerlendirilmiş; YZ destekli çift okumada kanser saptama oranı 6,7/1.000,

standart okumada 5,7/1.000 bulunmuş, geri çağırma oranı artmamıştır. Ancak radyologların YZ kullanmayı gönüllü seçmesi, randomize olmayan tasarım nedeniyle artık karıştırıcılık olasılığını gündeme getirir (Eisemann vd., 2025). Güney Kore'deki AI-STREAM prospektif kohortunda 24.543 kadında YZ desteği kanser saptama oranını 5,01/1.000'den 5,70/1.000'e yükseltmiş, geri çağırmada anlamlı değişiklik oluşturmamıştır; araştırmanın tek okuyuculu sisteme ve bir yıllık ilk sonuçlara dayandığı unutulmamalıdır (Chang vd., 2025). Danimarka'daki uygulama öncesi–sonrası değerlendirmede ise okuma yükü %33,5 azalmış ve tarama göstergelerinde iyileşme bildirilmiş, fakat dönemler arasındaki tarama aralığı farklılığı nedensel yorumu sınırlandırmıştır (Lauritzen vd., 2024).

DBT'de YZ, çok sayıda kesitin oluşturduğu okuma yükünü azaltma ve ince bulguları önceliklendirme açısından özellikle çekicidir. Retrospektif okuyucu çalışmalarında eş zamanlı YZ kullanımının doğruluk ve okuma süresini iyileştirebildiği; açıklama gereksinimi düşük etiketlerle eğitilen modellerin hem dijital mamografi hem DBT'de başarılı olabildiği gösterilmiştir (Conant vd., 2019; Lotter vd., 2021). Daha güncel prospektif eşleştirilmiş bir çalışmada YZ tabanlı triyaj ve karar desteği, mamografi/DBT taramasında insan okuma yükünü %63,6 azaltırken kanser saptama oranını 6,3/1.000'den 7,3/1.000'e çıkarmıştır (Elías-Cabot vd., 2026). Yine de DBT sonuçları iki boyutlu mamografiye otomatik olarak genellenmemeli; üretici, rekonstrüksiyon, sentetik mamografi ve kesit kalınlığı gibi değişkenler ayrıca doğrulanmalıdır.

Meme Yoğunluğu ve Ek Görüntüleme Kararı

YZ yazılımları meme yoğunluğunu sürekli bir ölçüm veya BI-RADS yoğunluk kategorisi olarak otomatikleştirebilir ve gözlemciler arası değişkenliği azaltabilir. Bununla birlikte yoğunluk, hem mamografik maskeleme hem de bağımsız kanser riskiyle ilişkili olmakla birlikte tek başına tamamlayıcı US ya da MRG endikasyonu değildir. Ek görüntüleme kararı; yaş, aile öyküsü, genetik yatkınlık, önceki biyopsiler ve hesaplanmış mutlak riskle birlikte verilmelidir. Ayrıca bir algoritmanın yoğun memede daha düşük ya da daha yüksek performans göstermesi model ve eşik değerine bağlıdır. Bu nedenle genel bir “yoğun memede YZ başarısı azalır” önermesi yerine, duyarlılık, özgüllük ve geri çağırma oranlarının yoğunluk kategorilerine göre raporlanması gerekir (Díaz vd., 2024; Gommers vd., 2026).

Tablo 1. Meme görüntüleme yapay zeka uygulamalarının klinik kanıt ve olgunluk düzeyi

Modalite / kullanım	Başlıca kanıt	Klinik olgunluk	Temel sınırlılık
2B tarama mamografisi	Randomize, prospektif ve gerçek yaşam çalışmaları	Yüksek; iş akışına bağlı	Sonuçlar tek bir yazılım veya tarama sistemine genellenemez
Dijital meme tomosentezi	Okuyucu çalışmaları ve prospektif eşleştirilmiş çalışma	Orta-yüksek	Kesit hacmi, üretici ve eşik değerine bağımlılık
Meme ultrasonografisi	Büyük retrospektif veri setleri ve okuyucu çalışmaları	Orta	Operatör bağımlılığı; gerçek yaşam biyopsi sonuçları sınırlı
Meme MRG	Küçük/orta ölçekli retrospektif çalışmalar ve derlemeler	Erken-orta	Protokol heterojenliği ve sınırlı prospektif doğrulama
Görüntü tabanlı risk tahmini	Çok merkezli retrospektif dış doğrulama	Araştırma / gelişen	Risk uyarlamalı taramanın klinik sonuç yararı henüz gösterilmiş değil

MEME ULTRASONOGRAFİSİNDE YAPAY ZEKA

Meme US'de YZ; lezyon saptama, segmentasyon, malignite olasılığı üretme ve iyi huylu-kötü huylu ayrımında karar desteği amacıyla kullanılmaktadır. Çoğu sistem BI-RADS kategorisini doğrudan belirlemekten çok, seçilmiş bir lezyona ya da tüm incelemeye ait malignite/abnormalite puanı üretir. Nihai BI-RADS değerlendirmesi; klinik bilgi, mamografi ve varsa önceki incelemelerle birlikte radyolog tarafından yapılır. Tek statik görüntüyle eğitilmiş bir modelin gerçek zamanlı tarama US'sine veya farklı cihazlardaki sine döngülerine uygulanması ayrıca doğrulama gerektirir.

Shen ve arkadaşlarının 288.767 US incelemesiyle geliştirdiği sistem, 44.755 incelemelik iç test setinde 0,976 AUROC elde etmiştir. Kanser açısından zenginleştirilmiş retrospektif okuyucu çalışmasında YZ desteği, on meme radyoloğunun yanlış pozitif değerlendirmelerini %37,3 ve simüle biyopsi önerilerini %27,8 azaltırken duyarlılığı korumuştur (Shen vd., 2021). Bu bulgu klinik potansiyeli güçlü biçimde destekler; ancak önerilen biyopsi sayısındaki simüle azalma, gerçek yaşamda biyopsi komplikasyonu veya kanser atlama oranının azaldığı anlamına gelmez. Tarama US'sinde operatörün lezyonu görüntüye dahil edip etmemesi, prob basısı, odak, kazanç, elastografi ve Doppler ayarları model performansını etkileyebilir. Bu

nedenle cihazlar arası ve operatörler arası dış doğrulama, mamografiye kıyasla daha kritik olabilir.

MEME MRG’SİNDE YAPAY ZEKA

Meme MRG’de YZ; otomatik meme ve lezyon segmentasyonu, hareket düzeltme, arka plan parankimal kontrastlanma ölçümü, kinetik analiz, malignite sınıflandırması ve neoadjuvan tedavi yanıtının tahmini için araştırılmaktadır. Çok fazlı dinamik kontrastlı seriler, T2 ağırlıklı görüntüler ve difüzyon verisinin birlikte işlenmesi tek sekanslı modellere göre daha zengin bilgi sunar; buna karşılık protokol, zaman çözünürlüğü, kontrast dozu ve üretici farklılıkları model aktarılabilirliğini güçleştirir. Mevcut literatürün önemli bölümü retrospektif, tek merkezli ve küçük örneklemlidir (Adam vd., 2023).

Jiang ve arkadaşlarının 111 kadın ve 19 meme radyoloğunu içeren retrospektif okuyucu çalışmasında YZ desteği ortalama AUC’yi 0,71’den 0,76’ya yükseltmiş; BI-RADS 3 eşliğinde duyarlılık %90’dan %94’e çıkmış, ancak özgüllükte anlamlı iyileşme gösterilmemiştir (Jiang vd., 2021). Bu sonuç, YZ’nin MRG yorumuna katkı sağlayabileceğini gösterir; fakat küçük bir okuyucu çalışması rutin klinik etkinlik kanıtı değildir. MRG’de tam otomatik lezyon saptama ve yüksek risk taramasının özerk biçimde yürütülmesi için çok merkezli prospektif çalışmalar, interval kanser takibi ve protokol bağımsızlığı gereklidir. Güncel durumda YZ’nin en olgun rolü, segmentasyon ve kantifikasyon gibi zaman alan yardımcı görevleri standardize etmek ve radyoloğun kararını desteklemektir.

GÖRÜNTÜ TABANLI RİSK TAHMİNİ VE KİŞİSELLEŞTİRİLMİŞ TARAMA

Geleneksel Gail, Tyrer–Cuzick ve benzeri modeller yaş, üreme öyküsü, aile öyküsü, genetik değişkenler ve meme yoğunluğu gibi risk etmenlerine dayanır. Görüntü tabanlı derin öğrenme modelleri ise mamografide insan gözüyle ayrı ayrı ölçülmeyen parankimal örüntülerden gelecekteki kanser riskini tahmin etmeyi amaçlar. Mirai modeli, farklı zaman ufuklarında risk üretmek ve eksik klinik değişkenlerden daha az etkilenmek üzere geliştirilmiş; ABD, İsveç ve Tayvan kohortlarında retrospektif olarak değerlendirilmiştir (Yala vd., 2021). Daha sonraki çok kurumlu doğrulama, modelin farklı merkezlerde umut verici ayırım gücünü koruduğunu göstermiştir (Yala vd., 2022). Sistematik incelemelerde görüntü tabanlı modellerin geleneksel risk araçlarından daha yüksek ayırım gücü bildirdiği

görülse de çalışmaların kalibrasyon, sonlanım ve takip süreleri heterojendir (Schopf vd., 2024).

Bu nedenle görüntü tabanlı risk puanı, bugün için ek MRG ya da US kararını tek başına veren bir test olarak değerlendirilmemelidir. Risk tahmininin yararı yalnızca yüksek AUC ile değil; modelin mutlak risk kalibrasyonu, farklı yaş ve etnik gruplardaki doğruluğu, önerilen tarama stratejisinin ileri evre ve interval kanserleri azaltıp azaltmadığı ve gereksiz inceleme yüküyle birlikte ölçülmelidir. Kişiselleştirilmiş tarama güçlü bir araştırma yönüdür; ancak mortalite veya ileri evre hastalık gibi hasta düzeyindeki sonuçlara katkısı prospektif olarak gösterilmeden standart tarama programlarının yerine geçirilmemelidir.

İNSAN-YAPAY ZEKA ETKİLEŞİMİ VE KLİNİK UYGULAMA

YZ ile radyoloğun birleşimi her zaman iki tarafın tek başına performansından daha iyi değildir. Algoritma doğru olduğunda karar desteği yararlı olabilir; yanlış olduğunda ise otomasyon yanlışlığı, yani kullanıcının otomatik öneriyi kendi değerlendirmesinin üzerinde tutması ortaya çıkabilir. Dratsch ve arkadaşlarının 27 radyologla yürüttüğü deneyde yanlış YZ önerileri deneyimsiz, orta deneyimli ve çok deneyimli okuyucuların tamamını etkilemiştir (Dratsch vd., 2023). Bu nedenle kullanıcı eğitiminin yalnızca arayüz kullanımını değil; eşik değerlerini, olası hata biçimlerini, yanlış güveni ve algoritma-radyolog uyumsuzluğunda izlenecek yolu kapsamaması gerekir.

Isı haritaları ve işaretlemeler hata analizini kolaylaştırabilir, ancak algoritmanın kararını nedensel olarak açıkladıklarını veya modelin güvenli olduğunu kanıtlamaz. Görsel açıklama, yanlış bölgeyi makul biçimde vurgulayabilir ya da model kararına sonradan üretilmiş bir gerekçe izlenimi verebilir. Açıklanabilirliğin yanında olasılık kalibrasyonu, belirsizlik bildirimi, önceki incelemelerle karşılaştırma, klinik bilgiye erişim ve kullanıcı davranışının izlenmesi gerekir (Ghassemi vd., 2021). En güvenli arayüz, radyoloğun bağımsız ilk değerlendirmesini koruyan ve YZ önerisini gerektiğinde ikinci aşamada gösteren tasarım olabilir; bunun etkisi yerel iş akışında ölçülmelidir.

Sorumluluk çerçevesi, sistemin rolüyle birlikte tanımlanmalıdır. Karar desteğinde nihai değerlendirme radyologda kalırken, bağımsız okuyucu veya insan okumasını dışlayan triyaj modelleri farklı bir risk ve sorumluluk dağılımı oluşturur. Hastalara YZ kullanımının nasıl bildirileceği, algoritma ile radyolog uyumsuzluğunun rapora yansıtılıp yansıtılmayacağı ve interval

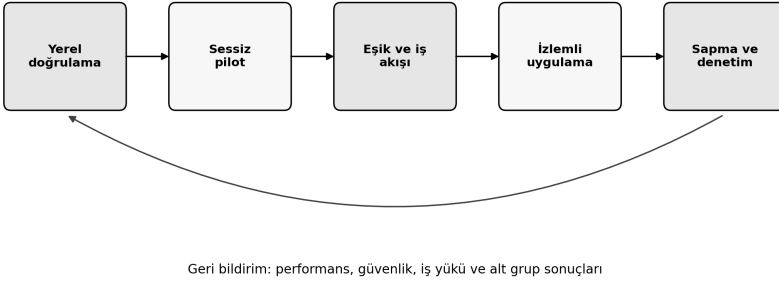
kanser incelemelerinde algoritma çıktılarının nasıl denetleneceği kurum politikalarında açık olmalıdır. Genel etik ve hukuki ilkeler kitabın ilgili bölümünde ele alınmakla birlikte, meme taramasına özgü kalite göstergeleri ayrıca tanımlanmalıdır.

SINIRLILIKLAR, YEREL DOĞRULAMA VE SÜREKLİ İZLEM

Bir YZ sisteminin geliştirme veri setindeki başarısı, kurumsal PACS ortamındaki performansı garanti etmez. Cihaz üreticisi, dedektör, görüntü işleme, kompresyon, tarama yaşı, kanser prevalansı ve meme yoğunluğu dağılımı değiştiğinde veri dağılımı kayabilir. Dış doğrulama mümkünse amaçlanan klinik popülasyonda, ardışık incelemeler üzerinde ve önceden belirlenmiş eşiklerle yapılmalıdır. Kanser olmayan mamogramlarda yeterli negatif referans için bir ya da iki yıllık kayıt bağlantısı ve takip gerekir; yalnızca kısa süre içinde biyopsi yapılmamasını “negatif” kabul etmek yanlış güven oluşturabilir.

Tek başına AUC klinik uygulama kararı için yeterli değildir. Tarama mamografisinde kanser saptama oranı, geri çağırma oranı, pozitif öngörü değeri, duyarlılık, özgüllük, interval kanser, ileri evre kanser ve okuma iş yükü birlikte raporlanmalıdır. US ve MRG’de lezyon düzeyi performansın yanında hasta düzeyi yanlış negatifler, biyopsi önerileri ve klinik yönetim değişikliği değerlendirilmelidir. Alt grup analizleri yaş, meme yoğunluğu, cihaz, merkez ve mümkün olduğunda demografik özelliklere göre önceden tanımlanmalı; düşük örneklemlili alt gruplarda kesinlik iddiasından kaçınılmalıdır.

Klinik entegrasyon aşamalı yürütülmelidir. Önce yerel retrospektif doğrulama yapılmalı, ardından sistemin klinik kararı etkilemediği sessiz pilot dönemde veri aktarımı, hata biçimleri ve uygun eşik gözlenmelidir. Eşik ve iş akışı tanımlandıktan sonra sınırlı kapsamlı, izlemlili kullanıma geçilmeli; kanser saptama, geri çağırma, yanlış negatifler, kullanıcı uyumu ve iş yükü düzenli denetlenmelidir. Yazılım güncellemesi, cihaz değişikliği veya popülasyon kayması olduğunda yeniden doğrulama tetiklenmelidir. Şekil 2, güvenli uygulamanın tek seferlik onaydan çok sürekli bir kalite döngüsü olduğunu özetlemektedir.



Şekil 2. Meme görüntülemeye yapay zekanın güvenli klinik uygulama döngüsü. Her yazılım güncellemesi, cihaz değişikliği veya anlamlı performans sapması yeniden doğrulama gerektirebilir.

SONUÇ VE ÖNERİLER

Meme radyolojisinde YZ için en güçlü klinik kanıt tarama mamografisinde birikmiştir. Randomize ve prospektif çalışmalar, doğru iş akışında YZ'nin okuma yükünü azaltırken kanser saptama ve duyarlılığı koruyabileceğini veya artırabileceğini göstermektedir. Bununla birlikte triyaj, eş zamanlı karar desteği, bağımsız okuyucu ve güvenlik ağı birbirinden farklı müdahalelerdir; bir modelde elde edilen yarar diğerine aktarılmalıdır. DBT'de kanıt gelişmektedir. US ve MRG'de sonuçlar umut verici olmakla birlikte, rutin özerk kullanım için prospektif gerçek yaşam verisi sınırlıdır. Görüntü tabanlı risk tahmini ise kişiselleştirilmiş tarama için güçlü bir araştırma alanı olmayı sürdürmektedir.

Klinik uygulamada ürün seçimi yalnızca yayımlanmış AUC değerine dayandırılmamalıdır. Sistem amaçlanan kullanım için yerel olarak doğrulanmalı; karar eşiği, kullanıcı arayüzü, arbitraj yolu ve sorumluluk açıkça tanımlanmalıdır. Performans meme yoğunluğu ve cihaz gibi alt gruplarda izlenmeli, otomasyon yanlılığını azaltacak eğitim verilmelidir. YZ radyoloğun dikkatini ve karar tutarlılığını güçlendirebilir; ancak güvenli katkı, teknolojinin klinik soruna uygun yerleştirilmesi ve sürekli kalite güvencesiyle mümkündür.

KAYNAKÇA

- Abu Abeelh, E., & Abuabeileh, Z. (2025). Screening mammography and artificial intelligence: A comprehensive systematic review. *Cureus*, 17(2), e79353. <https://doi.org/10.7759/cureus.79353>
- Adam, R., Dell'Aquila, K., Hodges, L., Maldjian, T., & Duong, T. Q. (2023). Deep learning applications to breast cancer detection by magnetic resonance imaging: A literature review. *Breast Cancer Research*, 25(1), 87. <https://doi.org/10.1186/s13058-023-01687-4>
- Chang, Y. W., Ryu, J. K., An, J. K., Choi, N., Park, Y. M., Ko, K. H., & Han, K. (2025). Artificial intelligence for breast cancer screening in mammography (AI-STREAM): Preliminary analysis of a prospective multicenter cohort study. *Nature Communications*, 16, 2248. <https://doi.org/10.1038/s41467-025-57469-3>
- Conant, E. F., Toledano, A. Y., Periaswamy, S., Fotin, S. V., Go, J., Boatsman, J. E., & Hoffmeister, J. W. (2019). Improving accuracy and efficiency with concurrent use of artificial intelligence for digital breast tomosynthesis. *Radiology: Artificial Intelligence*, 1(4), e180096. <https://doi.org/10.1148/ryai.2019180096>
- Dembrower, K., Crippa, A., Colón, E., Eklund, M., Strand, F., & ScreenTrustCAD Trial Consortium. (2023). Artificial intelligence for breast cancer detection in screening mammography in Sweden: A prospective, population-based, paired-reader, non-inferiority study. *The Lancet Digital Health*, 5(10), e703–e711. [https://doi.org/10.1016/S2589-7500\(23\)00153-X](https://doi.org/10.1016/S2589-7500(23)00153-X)
- Díaz, O., Rodríguez-Ruiz, A., & Sechopoulos, I. (2024). Artificial intelligence for breast cancer detection: Technology, challenges, and prospects. *European Journal of Radiology*, 175, 111457. <https://doi.org/10.1016/j.ejrad.2024.111457>
- Dratsch, T., Chen, X., Rezazade Mehrizi, M., Kloeckner, R., Mähringer-Kunz, A., Püsken, M., Baeßler, B., Sauer, S., Maintz, D., & Pinto dos Santos, D. (2023). Automation bias in mammography: The impact of artificial intelligence BI-RADS suggestions on reader performance. *Radiology*, 307(4), e222176. <https://doi.org/10.1148/radiol.222176>
- Eisemann, N., Bunk, S., Mukama, T., Baltus, H., Elsner, S. A., Gomille, T., Hecht, G., Heywang-Köbrunner, S., Rathmann, R., Siegmann-Luz, K., Töllner, T., Vomweg, T. W., Leibig, C., & Katalinic, A. (2025). Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. *Nature Medicine*, 31(3), 917–924. <https://doi.org/10.1038/s41591-024-03408-6>
- Eliás-Cabot, E., Romero-Martín, S., Raya-Povedano, J. L., Rodríguez-Ruiz, A., & Álvarez-Benito, M. (2026). AI-based triage and decision support in mammography and digital tomosynthesis for breast cancer screening: A paired, noninferiority trial. *Nature Medicine*, 32(4), 1296–1305. <https://doi.org/10.1038/s41591-026-04277-x>
- Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, 3(11), e745–e750. [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9)
- Gommers, J., Hernström, V., Josefsson, V., Sartor, H., Schmidt, D., Hjelmgren, A., Larsson, A.-M., Hofvind, S., Andersson, I., Rosso, A., Hagberg, O., & Lång, K. (2026). Interval cancer, sensitivity, and specificity comparing AI-supported mammography screening with standard double reading without AI in the MASAI study: A randomised, controlled, non-inferiority, single-blinded, population-based, screening-accuracy trial. *The Lancet*, 407(10527), 505–514. [https://doi.org/10.1016/S0140-6736\(25\)02464-X](https://doi.org/10.1016/S0140-6736(25)02464-X)
- Hernström, V., Josefsson, V., Sartor, H., Schmidt, D., Larsson, A.-M., Hofvind, S., Andersson, I., Rosso, A., Hagberg, O., & Lång, K. (2025). Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): A randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. *The Lancet Digital Health*, 7(3), e175–e183. [https://doi.org/10.1016/S2589-7500\(24\)00267-X](https://doi.org/10.1016/S2589-7500(24)00267-X)
- Jiang, Y., Edwards, A. V., & Newstead, G. M. (2021). Artificial intelligence applied to breast MRI for improved diagnosis. *Radiology*, 298(1), 38–46. <https://doi.org/10.1148/radiol.2020200292>
- Lång, K., Josefsson, V., Larsson, A.-M., Larsson, S., Högberg, C., Sartor, H., Hofvind, S., Andersson, I., & Rosso, A. (2023). Artificial intelligence-supported screen reading versus

- standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): A clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. *The Lancet Oncology*, 24(8), 936–944. [https://doi.org/10.1016/S1470-2045\(23\)00298-X](https://doi.org/10.1016/S1470-2045(23)00298-X)
- Lauritzen, A. D., Lillholm, M., Lynge, E., Nielsen, M., Karssemeijer, N., & Vejborg, I. (2024). Early indicators of the impact of using artificial intelligence in mammography screening for breast cancer. *Radiology*, 311(3), e232479. <https://doi.org/10.1148/radiol.232479>
- Lehman, C. D., Wellman, R. D., Buist, D. S. M., Kerlikowske, K., Tosteson, A. N. A., & Miglioretti, D. L. (2015). Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Internal Medicine*, 175(11), 1828–1837. <https://doi.org/10.1001/jamainternmed.2015.5231>
- Lotter, W., Diab, A. R., Haslam, B., Kim, J. G., Grisot, G., Wu, E., Wu, K., Onieva, J. O., Boxerman, J. L., Wang, M., Bandler, M., Vijayaraghavan, G., & Sorensen, A. G. (2021). Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nature Medicine*, 27(2), 244–249. <https://doi.org/10.1038/s41591-020-01174-9>
- McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafiyan, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., ... Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89–94. <https://doi.org/10.1038/s41586-019-1799-6>
- Ng, A. Y., Oberije, C. J. G., Ambrózay, É., Szabó, E., Serfőző, O., Karpati, E., Fox, G., Glocker, B., Morris, E. A., Forrai, G., & Kecskemethy, P. D. (2023). Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. *Nature Medicine*, 29(12), 3044–3049. <https://doi.org/10.1038/s41591-023-02625-9>
- Rodriguez-Ruiz, A., Lång, K., Gubern-Merida, A., Broeders, M., Gennaro, G., Clauser, P., Helbich, T. H., Chevalier, M., Tan, T., Mertelmeier, T., Wallis, M. G., Andersson, I., Zackrisson, S., Mann, R. M., & Sechopoulos, I. (2019). Stand-alone artificial intelligence for breast cancer detection in mammography: Comparison with 101 radiologists. *Journal of the National Cancer Institute*, 111(9), 916–922. <https://doi.org/10.1093/jnci/djy222>
- Schopf, C. M., Ramwala, O. A., Lowry, K. P., Hofvind, S., Marinovich, M. L., Houssami, N., Elmore, J. G., Dontchos, B. N., Lee, J. M., & Lee, C. I. (2024). Artificial intelligence-driven mammography-based future breast cancer risk prediction: A systematic review. *Journal of the American College of Radiology*, 21(2), 319–328. <https://doi.org/10.1016/j.jacr.2023.10.018>
- Shen, Y., Shamout, F. E., Oliver, J. R., Witowski, J., Kannan, K., Park, J., Wu, N., Huddleston, C., Wolfson, S., Millet, A., Ehrenpreis, R., Awal, D., Tyma, C., Samreen, N., Gao, Y., Chhor, C., Gandhi, S., Lee, C., Kumari-Subaiya, S., ... Geras, K. J. (2021). Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams. *Nature Communications*, 12, 5645. <https://doi.org/10.1038/s41467-021-26023-2>
- Yala, A., Mikhael, P. G., Strand, F., Lin, G., Smith, K., Wan, Y.-L., Lamb, L., Hughes, K., Lehman, C. D., & Barzilay, R. (2021). Toward robust mammography-based models for breast cancer risk. *Science Translational Medicine*, 13(578), eaba4373. <https://doi.org/10.1126/scitranslmed.aba4373>
- Yala, A., Mikhael, P. G., Strand, F., Lin, G., Satuluru, S., Kim, T., Banerjee, I., Gichoya, J., Trivedi, H., Lehman, C. D., Hughes, K., Sheedy, D. J., Matthis, L. M., Karunakaran, B., Hegarty, K. E., Sabino, S., Silva, T. B., Evangelista, M. C., Caron, R. F., ... Barzilay, R. (2022). Multi-institutional validation of a mammography-based breast cancer risk model. *Journal of Clinical Oncology*, 40(16), 1732–1740. <https://doi.org/10.1200/JCO.21.01337>

ABDOMEN RADYOLOJİSİNDE YAPAY ZEKA: ORGAN SEGMENTASYONU VE LEZYON KARAKTERİZASYONU

Mehmet Ali Özel*

GİRİŞ

Abdominal radyoloji; çok sayıda organı, farklı doku kontrastlarını, değişken çekim protokollerini ve geniş bir hastalık yelpazesini aynı inceleme içinde değerlendirir. Bilgisayarlı tomografi (BT) ve manyetik rezonans görüntüleme (MRG) veri hacmindeki artış, küçük lezyonların saptanması, çok fazlı incelemelerin karşılaştırılması ve nicel ölçümlerin tekrarlanabilir biçimde yapılması gereksinimini belirginleştirmiştir. Yapay zeka (YZ), bu bağlamda radyoloğun yerine geçen tek bir tanı aracı değil; segmentasyon, ölçüm, saptama, karakterizasyon, triyaj ve raporlama gibi birbirinden farklı görevlerde kullanılabilen bir yöntemler bütünüdür (Hosny vd., 2018; Sahiner vd., 2019).

Teknik performans ile klinik yarar aynı kavram değildir. Bir model seçilmiş bir test kümesinde yüksek Dice katsayısına veya eğri altında kalan alana (AUC) ulaşabilir; buna rağmen farklı cihaz, kontrast fazı, rekonstrüksiyon çekirdeği ya da hasta popülasyonunda başarısız olabilir. Abdominal uygulamalarda asıl soru yalnızca algoritmanın doğru çalışıp çalışmadığı değil, çıktının radyoloğun kararını, rapor tutarlılığını, işlem süresini ve hasta yönetimini güvenli biçimde iyileştirip iyileştirmediğidir. Bu nedenle her kullanım alanı, amaçlanan klinik görev ve uygun sonlanım ölçütleri üzerinden değerlendirilmelidir.

Bu bölüm; abdominal BT ve MRG’de organ segmentasyonu, fırsatçı tarama, lezyon saptama ve karakterizasyon, akut abdomen triyajı, radyomik ve çok modlu modeller ile üretken YZ destekli raporlamayı ele almaktadır. Tedavi yanıtı ve onkolojik prognoz gibi konular, kitabın ilgili bölümleriyle yinelemeyi önlemek amacıyla yalnızca gerektiği ölçüde anılmaktadır. Bölüm boyunca araştırma sonuçları ile rutin klinik uygulama birbirinden ayrılmış;

* Dr. Öğretim Üyesi, Düzce Üniversitesi, Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, drmaliozel@gmail.com, ORCID: <https://orcid.org/0000-0001-8817-5769>

dış doğrulama, kalibrasyon ve sürekli performans izlemi özellikle vurgulanmıştır.

ABDOMİNAL GÖRÜNTÜLEMEDE SEGMENTASYON

Teknik Çerçeve ve Performans Ölçütleri

Segmentasyon, görüntüdeki her vokselin karaciğer, pankreas, böbrek, dalak, damar, kas veya lezyon gibi bir anatomik sınıfa atanmasıdır. Çıktı; organ hacmi, lezyon yükü, cerrahi planlama, vücut bileşimi ve radyomik analiz için temel oluşturur. U-Net ailesi ve kendi kendini yapılandıran nnU-Net, tıbbi görüntü segmentasyonunda güçlü başvuru yaklaşımlarıdır. Dönüştürücü tabanlı mimariler geniş bağlamı modelleyebilse de belirli bir mimariyi evrensel “altın standart” kabul etmek doğru değildir; veri boyutu, ön işleme, kayıp fonksiyonu ve görev tanımı sonucu üzerinde en az mimari kadar etkilidir (Isensee vd., 2021; Ram vd., 2025).

En sık bildirilen örtüşme ölçütü Dice katsayısıdır; 0 ile 1 arasında değişir ve tahmin edilen maske ile referans maskenin ortaklığını gösterir. Bununla birlikte yüksek Dice, özellikle küçük lezyonlarda klinik olarak önemli bir sınır hatasını gizleyebilir. Bu nedenle ortalama yüzey mesafesi veya Hausdorff mesafesi gibi sınır ölçütleri, hacim farkı ve tamamen başarısız segmentasyon oranı da raporlanmalıdır. Klinik amaç çap ölçümü ise ölçüm yanlılığı ve gözlemciler arası değişkenlik; işlem planlaması ise kritik komşuluklarda sınır doğruluğu ayrıca değerlendirilmelidir. Model çıktısının görsel kalite kontrolü yapılmadan otomatik hacim değerinin rapora aktarılması güvenli değildir.

Performans; kesit kalınlığı, kontrast zamanı, artefakt, anatomik varyasyon, geçirilmiş cerrahi ve yaygın hastalık gibi değişkenlerle katmanlandırılmalıdır. Örneğin normal bir pankreasın segmentasyonu ile atrofik, ödemli veya tümörle konturu bozulmuş pankreasın segmentasyonu aynı güçlükte değildir. Ortalama metrikler, nadir fakat klinik açıdan kritik hata kümelerini maskeleyebilir. Bu nedenle örnek bazlı kalite bayrağı, belirsizlik tahmini ve insan düzeltme süresi, salt ortalama Dice değerinden daha işlevsel olabilir.

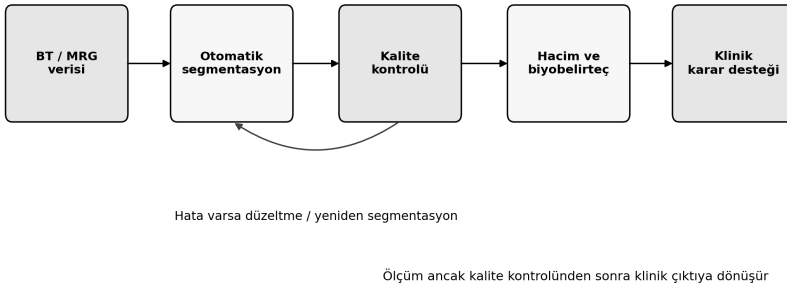
Çoklu Organ Segmentasyonu

Çoklu organ modelleri tek bir incelemede onlarca anatomik yapıyı eş zamanlı olarak işaretleyebilir. TotalSegmentator’ın özgün BT çalışması 104 anatomik yapıyı kapsayan bir model sunmuş; açık veri ve kod yaklaşımı nedeniyle araştırma ve iş akışı geliştirmede yaygınlaşmıştır (Wasserthal vd.,

2023). Daha sonra genişletilen veri kümelerinde yapı sayısı artmış olsa da özgün makaledeki sayı 104'tür. MRG için geliştirilen TotalSegmentator MRI ise sekanslardan daha bağımsız çalışmayı hedefleyen ve 80 yapıyı segmentleyen ayrı bir modeldir; BT modeliyle aynı ürün ya da aynı doğrulama problemi olarak görülmemelidir (Akinci D'Antonoli vd., 2025).

Genel amaçlı modeller; radyoterapi planlama, cerrahi hazırlık, organ hacimlerinin izlenmesi ve görüntü arşivlerinin otomatik etiketlenmesi için başlangıç noktası sağlayabilir. Buna karşılık açık veri setindeki başarı, yerel klinik popülasyona doğrudan aktarılamaz. Kontrastlı ve kontrastsız seriler, portal venöz ve arteriyel fazlar, düşük doz protokoller ve MRG sekansları ayrı doğrulama gerektirir. Küçük damarlar, ince barsak duvarı ve patoloji nedeniyle sınırı silinmiş organlarda manuel düzeltme gereksinimi artabilir.

Tıbbi görüntüler için uyarlanan MedSAM, genel görüntü segmentasyonu için geliştirilen Segment Anything Model'in tıbbi verilerle ince ayar yapılmış, istem tabanlı bir sürümüdür. Nokta veya kutu gibi kullanıcı istemleriyle farklı organ ve lezyonların işaretlenmesini kolaylaştırabilir; ancak nadir patolojiyi klinik bağlam olmadan kendiliğinden tanıyan özerk bir sistem değildir. İstem kalitesi, görüntü düzlemi ve hedefin sınır belirginliği çıktıyı etkiler (Ma vd., 2024). Bu nedenle temel modellerin en gerçekçi yakın dönem rolü, insanın başlattığı segmentasyonu hızlandırmak ve elle düzeltme yükünü azaltmaktır.



Şekil 1. Abdominal görüntülemede segmentasyondan klinik çıktıya uzanan süreç. Otomatik maske, görsel kalite kontrolünden geçmeden hacim veya biyobelirteç olarak kullanılmamalıdır.

FIRSATÇI TARAMA VE OTOMATİK VÜCUT BİLEŞİMİ

Başka bir klinik endikasyonla çekilmiş abdominal BT'den ek risk göstergeleri elde edilmesi fırsatçı tarama olarak adlandırılır. Vertebra kemik yoğunluğu, abdominal aort kalsifikasyonu, visseral ve subkutan yağ alanı, iskelet kası miktarı ve karaciğer yağlanması aynı incelemeden ölçülebilir. YZ, organ ve doku segmentasyonunu otomatikleştirerek bu ölçümlerin yüksek hacimli iş akışlarında uygulanabilirliğini artırır. Büyük retrospektif kohortlarda otomatik BT biyobelirteçlerinin kardiyovasküler olaylar, metabolik risk ve uzun dönem olumsuz sonuçlarla ilişkili olduğu gösterilmiştir (O'Connor vd., 2019; Pickhardt vd., 2020; Lee vd., 2023).

Vücut bileşimi ölçümleri geleneksel olarak belirli bir lomber düzeyde tek kesit üzerinden yapılırken, üç boyutlu segmentasyon tüm hacmi değerlendirebilir. Kas alanı veya kas dansitesi, özellikle yaşlı ve kronik hastalığı olan bireylerde kırılabilirlik ve sarkopeniyle ilişkili bilgi sağlayabilir. Visseral yağ ve karaciğer yağlanması kardiyometabolik riskin görünür olmayan bileşenlerini ortaya çıkarabilir. Bununla birlikte görüntü tabanlı bir eşik, klinik tanı yerine geçmez; yaş, cinsiyet, vücut yapısı, kontrast kullanımı ve popülasyon özelliklerine göre kalibre edilmelidir (Pickhardt vd., 2021; Pickhardt, 2022).

Fırsatçı taramanın yararı, yalnızca otomatik ölçüm üretmekle sınırlı değildir. Ölçümün hangi eşikte raporlanacağı, kime ve nasıl bildirileceği, doğrulama incelemesi gerektirip gerektirmediği ve takip sorumluluğunun kimde olduğu tanımlanmalıdır. Çok sayıda düşük özgüllüklü uyarı, klinisyenlerde alarm yorgunluğu ve gereksiz tetkik oluşturabilir. Bu nedenle kurumlar teknik doğruluğun yanı sıra bildirim eşiği, yönlendirme yolu, hasta düzeyindeki yarar ve ek iş yükünü birlikte değerlendirmelidir.

LEZYON SAPTAMA VE KARAKTERİZASYONU

Lezyon saptama, görüntüde şüpheli bir odağın bulunmasını; karakterizasyon ise bu odağın iyi huylu-kötü huylu ayrımı, alt tip veya risk kategorisi açısından değerlendirilmesini ifade eder. Bu iki görev aynı değildir. Bir model lezyonu doğru bulabilir fakat yanlış sınıflandırabilir; başka bir model ise yalnızca radyoloğun seçtiği lezyon üzerinde çalışabilir. Sonuçlar hasta düzeyi ve lezyon düzeyi ayrı raporlanmalı; duyarlılığın yanında inceleme başına yanlış pozitif sayısı ve kalibrasyon da verilmelidir.

Karaciğer

Karaciğer YZ uygulamaları; fokal lezyon saptama, segmentasyon, hacim ölçümü, yağlanma ve fibrozis göstergeleri ile hepatoselüler karsinom risk sınıflamasını kapsar. Dinamik kontrastlı BT ve MRG’de arteriyel faz kontrastlanması, yıkanma görünümü, kapsül ve büyüme gibi özelliklerin değerlendirilmesi gerekir. LI-RADS 2018 terminolojisinde başlıca özellikler nonrim arteriyel faz hiper-kontrastlanması, nonperiferik yıkanma, kontrastlanan kapsül görünümü ve eşik büyümedir; algoritma çalışmaları bu standart tanımlarla uyumlu olmalıdır (Chernyak vd., 2018).

Derin öğrenme ile LI-RADS kategori tahmini ve dinamik BT’de karaciğer kitlelerinin ayrımı üzerine çalışmalar umut verici sonuçlar bildirmiştir (Yasaka vd., 2018; Yamashita vd., 2020). Bununla birlikte mevcut kanıtın önemli bölümü retrospektif, seçilmiş ve lezyon açısından zenginleştirilmiş veri setlerine dayanır. LI-RADS kategorisi yalnızca görüntü örüntüsünden ibaret değildir; risk popülasyonu, önceki inceleme, tedavi öyküsü ve görüntü kalitesi de değerlendirmeye katılır. Bu nedenle kategori öneren bir model, eksik klinik bağlamı görünür biçimde bildirmeli ve radyolog kararının yerine geçmemelidir.

Karaciğer lezyonlarında segmentasyon; tümör yükü, ablasyon planlama ve longitudinal ölçüm için yararlı olabilir. Ancak düzensiz sınırlı, infiltratif veya tedavi edilmiş lezyonlarda referans maskenin kendisi gözlemciler arasında değişebilir. Model performansı değerlendirilirken yalnızca maske örtüşmesi değil, klinik karar açısından kritik çap, damar ilişkisi ve canlı tümör bileşeni gibi ölçümler de göz önüne alınmalıdır.

Pankreas

Pankreas duktal adenokarsinomu küçük boyutta silik olabilir; kontur değişikliği, kanal kesilmesi, distal atrofi ve damar ilişkisi gibi dolaylı bulgular tanıda önem taşır. Pankreasın ince ve değişken anatomisi, otomatik segmentasyonu karaciğer veya dalaktan daha güç hale getirir. Derin öğrenme tabanlı sistemler kontrastsız BT’de dahi pankreas kanseri saptama potansiyeli göstermiştir. PANDA çalışması, çok merkezli geniş veri üzerinde nonkontrast BT’den pankreas kanseri saptamaya yönelik dikkat çekici sonuçlar bildirmiştir (Cao vd., 2023).

Bu bulgu tarama veya erken tanı açısından umut vericidir; ancak modelin mikroskobik infiltrasyonu gördüğü anlamına gelmez. Performans; belirgin kitle, dolaylı bulgu, tümör evresi ve klinik semptomlara göre değişebilir.

Pankreatit, yağlı replasman, cerrahi değişiklik ve kistik lezyonlar yanlış pozitif veya yanlış negatif sonuçlara yol açabilir. Rutin uygulama için ardışık klinik popülasyonda dış doğrulama, inceleme başına uyarı yükü ve gecikmiş tanı üzerindeki etkinin prospektif değerlendirilmesi gerekir.

Böbrek ve Adrenal Bezler

Böbrek uygulamalarında otomatik parankim ve tümör segmentasyonu; lezyon hacmi, nefron koruyucu cerrahi planlama ve seri inceleme karşılaştırmasını destekleyebilir. Çok fazlı BT ve multiparametrik MRG radyomikleri, berrak hücreli renal hücreli karsinomda Dünya Sağlık Örgütü/Uluslararası Ürolojik Patoloji Derneği (WHO/ISUP) derecesini öngörmek için araştırılmıştır (Cui vd., 2020). Ancak histolojik derece tahmini güncel durumda araştırma aracıdır; eski Fuhrman derecelendirmesi yerine WHO/ISUP standardı kullanılmalı ve görüntü tahmini patoloji sonucunun eş değeri olarak sunulmamalıdır.

Adrenal nodüllerde temel klinik görev; kontrastsız dansite, kontrast yıkanma özellikleri, kimyasal kayma MRG bulguları, boyut ve onkolojik bağlamın birlikte yorumlanmasıdır. YZ bu ölçümleri otomatikleştirebilir ve önceki incelemelerdeki nodülü eşleştirebilir. Bununla birlikte hormon aktivitesi görüntüden doğrudan belirlenemez. Model çıktısı klinik ve biyokimyasal değerlendirmeyi tamamlayan bir karar desteği olarak konumlandırılmalıdır.

Gastrointestinal Sistem

Gastrointestinal sistemde YZ; mide ve kolorektal tümörlerin saptanması, BT kolonografide polip işaretleme, barsak duvar kalınlığı ve inflamasyon ölçümü gibi görevlerde araştırılmaktadır. Nonkontrast BT ile mide kanseri taramasına yönelik büyük ölçekli bir 2025 çalışması, hastalık saptama potansiyelini geniş bir veri kümesinde göstermiştir (Hu vd., 2025). BT kolonografide derin öğrenme tabanlı bilgisayar destekli saptama sistemleri kolorektal neoplaziler için geliştirilmektedir (Endo vd., 2025).

Bu alanlarda barsak distansiyonu, peristaltizm, dışkı, metal artefaktı ve inflamatuvar değişiklikler model performansını belirgin biçimde etkileyebilir. Kanser açısından zenginleştirilmiş veri setleri gerçek tarama prevalansını yansıtmayabilir; küçük özgüllük kayıpları bile çok sayıda yanlış pozitif uyarı oluşturabilir. Ayrıca görüntü tabanlı risk puanı, endoskopi ve histopatoloji gibi referans standartların yerini almamalıdır.

Tablo 1. Abdominal görüntülemelerde başlıca yapay zeka görevleri ve değerlendirme ölçütleri

Görev	Klinik çıktı ve uygun ölçüt	Mevcut kanıt	Başlıca sınırlılık
Organ / lezyon segmentasyonu	Hacim, çap ve harita; Dice, yüzey mesafesi, başarısız olgu oranı	Çok merkezli teknik doğrulamalar mevcut	Küçük lezyonlar, sınır belirsizliği ve protokol değişimi
Fırsatçı tarama	Kas, yağ, kemik ve kalsifikasyon ölçümü; kalibrasyon ve klinik sonlanım	Büyük retrospektif kohortlar mevcut	Eşiklerin popülasyon ve protokole bağımlılığı
Lezyon saptama	Hasta ve lezyon düzeyi duyarlılık; inceleme başına yanlış pozitif	Organ ve hastalığa göre değişken	Seçilmiş veri setleri ve spektrum yanlışlığı
Karakterizasyon / risk	AUC, duyarlılık, özgüllük, kalibrasyon ve karar eğrisi	Çoğunlukla retrospektif	Referans standart ve dış doğrulama eksikliği
Akut triyaj	Gecikme süresi, duyarlılık ve iş yükü	Erken uygulama ve retrospektif çalışmalar	Nadir kritik bulgular ve yanlış güven
Üretken raporlama	Doğruluk, eksiklik, tutarlılık ve insan düzeltme yükü	Erken dönem okuyucu / uygulanabilirlik çalışmaları	Olgusal hata, atlama ve otomasyon yanlışlığı

AKUT ABDOMEN VE ACİL İŞ AKIŞI

Acil abdominal görüntülemelerde amaç yalnızca tanısal doğruluk değil, kritik bulgunun zamanında görünür hale getirilmesidir. Serbest intraperitoneal hava, aktif kanama, aort patolojileri, barsak obstrüksiyonu, apandisit veya iskemi şüphesi için geliştirilen modeller iş listesini önceliklendirebilir ve radyoloğa olası bulgu bölgesini gösterebilir. Bu sistemlerin klinik değeri; hasta düzeyi duyarlılık, inceleme başına yanlış pozitif, raporlama gecikmesi ve iş yükü üzerindeki etkiyle ölçülmelidir.

Akut abdomen modellerinin önemli bölümü retrospektif veriyle geliştirilmiştir. Nadir fakat kritik olgularda eğitim verisinin azlığı, kontrast fazı ve protokol değişkenliği, birden çok eş zamanlı patoloji ve ameliyat sonrası anatomi temel güçlüklerdir. Yüksek öncelik puanı tanıyı doğrulamaz; düşük puan da incelemenin güvenle ertelenebileceği anlamına gelmez. Bu nedenle triyaj sistemleri, tanı koyan özerk araçlar yerine zaman yönetimini destekleyen ve başarısız olduğunda standart okuma akışını bozmayan güvenlik katmanları olarak tasarlanmalıdır.

RADYOMİK, ÇOK MODLU MODELLER VE TEMEL MODELLER

Radyomik; bir görüntü veya tanımlanmış ilgi hacminden yoğunluk, şekil ve doku gibi çok sayıda nicel özelliğin çıkarılmasıdır. Bu özellikler görüntüdeki makroskopik örüntüleri tanımlar; mikroskopik tümör heterojenliğini doğrudan ölçmez. Tekrarlanabilirlik; segmentasyon, vokselle boyutu, gri düzey ayrıklaştırması, rekonstrüksiyon, kontrast fazı ve özellik tanımlarına bağlıdır. Görüntü Biyobelirteç Standardizasyon Girişimi (IBSI), özellik hesaplamalarını ve raporlamayı uyumlaştırmak için temel bir çerçeve sunmaktadır (Zwanenburg vd., 2020).

Çok modlu modeller görüntüyü yaş, laboratuvar, klinik öykü, patoloji veya genomik verilerle birleştirebilir. Bu yaklaşım, yalnız görüntü kullanan bir modelden farklıdır. Görüntü-dil modeli ise görsel özelliklerle metinsel temsilleri ortak bir yapıda işler; her çok modlu model görüntü-dil modeli değildir. Klinik değişkenlerin eksikliği, ölçüm zamanlarının uyumsuzluğu ve kurumlar arası kodlama farkları, çok modlu modellerin aktarılabilirliğini sınırlandırabilir.

Temel modeller, geniş ve çeşitli veriler üzerinde ön eğitim alıp farklı görevlere uyarlanmayı amaçlar. MedSAM gibi istem tabanlı segmentasyon modelleri elle çizimi hızlandırabilir; görüntü-dil modelleri bulgu özetleme veya rapor taslağı oluşturma potansiyeli taşır. Ancak genel amaçlı bir modelin abdominal radyolojiye uyarlanması, yerel terminoloji ve klinik bağlamla doğrulama gereksinimini ortadan kaldırmaz. Özellikle çok nadir hastalıklar, ameliyat sonrası anatomi ve beklenmedik kombinasyonlar, eğitim dağılımı dışı hata riskini artırır.

RAPORLAMA OTOMASYONU VE ÜRETKEN YAPAY ZEKA

Üretken YZ, yapılandırılmış bulguları akıcı metne dönüştürme, raporu sadeleştirme, önceki incelemelerle değişimi özetleme ve hasta odaklı açıklama üretme amacıyla kullanılabilir. Üretken modellerle hasta merkezli radyoloji raporları üzerine yapılan çalışmalar, anlaşılabilirliği artırma potansiyelini göstermektedir (Park vd., 2024). Bununla birlikte GPT-4o, Med-PaLM 2 veya benzeri genel modeller, yalnızca adları veya genel yetenekleri nedeniyle onaylı, özerk abdominal raporlama sistemi olarak kabul edilemez.

Üretken model hatası; görüntü saptama sistemindeki yanlış pozitiften farklıdır. Model, raporda bulunmayan bir bulguyu ekleyebilir, var olan

bulguyu atlayabilir, laterality veya ölçüm değiştirebilir ve ikna edici fakat olgusal olarak yanlış bir gerekçe üretebilir. Bu durum “halüsinasyon” olarak adlandırılır. Güvenli kullanımda model yalnızca doğrulanmış girdilerle sınırlandırılmalı, kaynak bulgular izlenebilir olmalı, kritik ölçümler otomatik karşılaştırılmalı ve nihai rapor radyolog tarafından onaylanmalıdır. İnsan düzeltme süresi ile hata türleri, hız kazanımı kadar önemli sonuçlandırmalarıdır.

KLİNİK DOĞRULAMA, SINIRLILIKLAR VE SÜREKLİ İZLEM

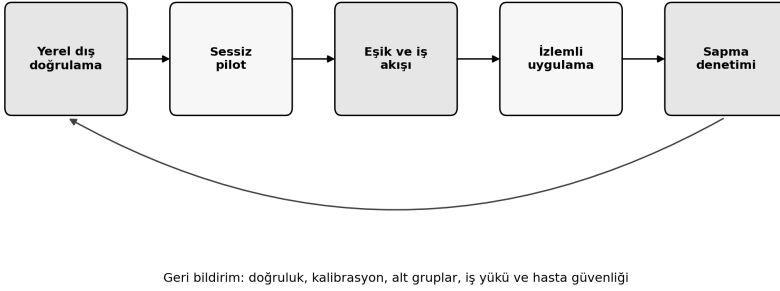
Bir modelin yayınlanmış performansı, satın alma veya klinik devreye alma için tek başına yeterli değildir. Amaçlanan kullanım; hedef popülasyon, modalite, kontrast fazı, cihazlar, klinik görev, kullanıcı ve karar eşiğiyle birlikte yazılı olarak tanımlanmalıdır. Çalışmaların raporlanmasında CLAIM, erken klinik değerlendirmede ise DECIDE-AI gibi çerçeveler şeffaflık sağlar (Mongan vd., 2020; Vasey vd., 2022). Yerel doğrulamada ardışık olgular, önceden belirlenmiş eşik ve klinik olarak anlamlı referans standart kullanılmalıdır.

AUC, farklı eşiklerdeki ayırma gücünü özetler; ancak prevalans, kalibrasyon ve klinik sonuç hakkında tek başına bilgi vermez. Duyarlılık ve özgüllüğün yanında pozitif ve negatif öngörü değerleri, kalibrasyon eğrisi, karar eğrisi, inceleme başına yanlış pozitif ve başarısız işleme oranı raporlanmalıdır. Segmentasyonda Dice ve yüzey mesafesine ek olarak manuel düzeltme süresi; triyajda raporlama gecikmesi; raporlamada olgusal hata ve kullanıcı düzeltme yükü ölçülmelidir.

Protokol, cihaz, rekonstrüksiyon yazılımı, kontrast madde ve hasta profili zaman içinde değişebilir. Bu veri dağılımı kayması performansı sessizce düşürebilir. İzlem; genel ortalamayla sınırlı kalmamalı, yaş, cinsiyet, cihaz, merkez, kontrast fazı ve klinik alt gruplara göre yapılmalıdır. Isı haritaları yararlı bir hata inceleme aracı olabilir; ancak kararın nedensel açıklaması veya model güvenliğinin kanıtı değildir (Ghassemi vd., 2021). Belirsizlik ve başarısızlık işaretleri, kullanıcıya anlaşılır biçimde gösterilmelidir.

Federe öğrenme, kurumların ham hasta verisini merkezi bir depoda paylaşmadan ortak model eğitmesine olanak verebilir. Bu yaklaşım veri hareketini azaltır; fakat gizliliği, veri kalitesini veya adaleti kendiliğinden garanti etmez. Model güncellemelerinden bilgi sızması, kurumlar arası veri dengesizliği, güvenlik ve yönetim sorunları ayrıca ele alınmalıdır (Sheller vd., 2020). Dolayısıyla federe öğrenme etik sorunun tek başına çözümü değil, uygun teknik ve hukuki önlemlerle kullanılabilecek bir iş birliği yöntemidir.

Klinik entegrasyon aşamalı olmalıdır. Önce yerel dış doğrulama yapılmalı; ardından sistemin kararı etkilemediği sessiz pilot dönemde veri akışı, gecikme ve hata örüntüleri gözlenmelidir. Eşik ve sorumluluk yolu belirlendikten sonra sınırlı, izlemli kullanıma geçilmeli; yazılım güncellemesi, cihaz değişimi veya anlamlı performans sapması yeniden doğrulamayı tetiklemelidir. Şekil 2, klinik uygulamanın tek seferlik onay değil, sürekli bir kalite döngüsü olduğunu göstermektedir.



Şekil 2. Abdominal radyolojide yapay zekanın güvenli klinik uygulama döngüsü. Performans sapması, yazılım güncellemesi veya cihaz değişikliği yeniden doğrulama gerektirebilir.

SONUÇ VE ÖNERİLER

Abdominal radyolojide YZ'nin en olgun katkısı segmentasyon ve fırsatçı taramadaki nicel ölçümlerdir. Organ bazlı lezyon saptama ve karakterizasyon araştırmaları umut verici olsa da kanıt çoğunlukla retrospektiftir ve farklı protokollere doğrudan genellenemez. Ürün seçimi yalnızca Dice veya AUC değerine dayanmamalı; amaçlanan kullanım, yerel doğrulama, kalite kontrol ve başarısızlık yolu tanımlanmalı, üretken modeller radyolog gözetiminde kullanılmalıdır.

KAYNAKÇA

- Akinci D'Antonoli, T., Berger, L. K., Indrakanti, A. K., Vishwanathan, N., Weiss, J., Jung, M., Berkarda, Z., Rau, A., Reiser, M., Küstner, T., Walter, A., Merkle, E. M., Boll, D. T., Breit, H.-C., Nicoli, A. P., Segeroth, M., Cyriac, J., Yang, S., & Wasserthal, J. (2025). TotalSegmentator MRI: Robust sequence-independent segmentation of multiple anatomic structures in MRI. *Radiology*, 314(2), e241613. <https://doi.org/10.1148/radiol.241613>
- Cao, K., Xia, Y., Yao, J., Han, X., Lambert, L., Zhang, T., Tang, W., Jin, G., Jiang, H., Fang, X., Nogues, I., Li, X., Guo, W., Wang, Y., Fang, W., Qiu, M., Hou, Y., Kovarnik, T., Vocka, M., ... Lu, J. (2023). Large-scale pancreatic cancer detection via non-contrast CT and deep learning. *Nature Medicine*, 29(12), 3033–3043. <https://doi.org/10.1038/s41591-023-02640-w>
- Chernyak, V., Fowler, K. J., Kamaya, A., Kielar, A. Z., Elsayes, K. M., Bashir, M. R., Kono, Y., Do, R. K., Mitchell, D. G., Singal, A. G., Tang, A., & Sirlin, C. B. (2018). Liver Imaging Reporting and Data System (LI-RADS) Version 2018: Imaging of hepatocellular carcinoma in at-risk patients. *Radiology*, 289(3), 816–830. <https://doi.org/10.1148/radiol.2018181494>
- Cui, E., Li, Z., Ma, C., Li, Q., Lei, Y., Lan, Y., Yu, J., Zhou, Z., Li, R., Long, W., & Lin, F. (2020). Predicting the ISUP grade of clear cell renal cell carcinoma with multiparametric MR and multiphase CT radiomics. *European Radiology*, 30(5), 2912–2921. <https://doi.org/10.1007/s00330-019-06601-1>
- Endo, S., Nagata, K., Utano, K., Nozu, S., Yasuda, T., Takabayashi, K., Hirayama, M., Togashi, K., & Ohira, H. (2025). Development and validation of computer-aided detection for colorectal neoplasms using deep learning incorporated with computed tomography colonography. *BMC Gastroenterology*, 25, 149. <https://doi.org/10.1186/s12876-025-03742-0>
- Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, 3(11), e745–e750. [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9)
- Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, 18(8), 500–510. <https://doi.org/10.1038/s41568-018-0016-5>
- Hu, C., Xia, Y., Zheng, Z., Cao, M., Zheng, G., Chen, S., Sun, J., Chen, W., Zheng, Q., Pan, S., Zhang, Y., Chen, J., Yu, P., Xu, J., Xu, J., Qiu, Z., Lin, T., Yun, B., Yao, J., ... Cheng, X. (2025). AI-based large-scale screening of gastric cancer from noncontrast CT imaging. *Nature Medicine*, 31(9), 3011–3019. <https://doi.org/10.1038/s41591-025-03785-6>
- Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2), 203–211. <https://doi.org/10.1038/s41592-020-01008-z>
- Lee, M. H., Zea, R., Garrett, J. W., Graffy, P. M., Summers, R. M., & Pickhardt, P. J. (2023). Abdominal CT body composition thresholds using automated AI tools for predicting 10-year adverse outcomes. *Radiology*, 306(2), e220574. <https://doi.org/10.1148/radiol.220574>
- Ma, J., He, Y., Li, F., Han, L., You, C., & Wang, B. (2024). Segment anything in medical images. *Nature Communications*, 15, 654. <https://doi.org/10.1038/s41467-024-44824-z>
- Mongan, J., Moy, L., & Kahn, C. E. (2020). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A guide for authors and reviewers. *Radiology: Artificial Intelligence*, 2(2), e200029. <https://doi.org/10.1148/ryai.2020200029>
- O'Connor, S. D., Graffy, P. M., Zea, R., & Pickhardt, P. J. (2019). Does nonenhanced CT-based quantification of abdominal aortic calcification outperform the Framingham Risk Score in predicting cardiovascular events in asymptomatic adults? *Radiology*, 290(1), 108–115. <https://doi.org/10.1148/radiol.2018180562>
- Park, J., Oh, K., Han, K., & Lee, Y. H. (2024). Patient-centered radiology reports with generative artificial intelligence: Adding value to radiology reporting. *Scientific Reports*, 14, 13218. <https://doi.org/10.1038/s41598-024-63824-z>
- Pickhardt, P. J. (2022). Value-added opportunistic CT screening: State of the art. *Radiology*, 303(2), 241–254. <https://doi.org/10.1148/radiol.211561>
- Pickhardt, P. J., Graffy, P. M., Perez, A. A., Lubner, M. G., Elton, D. C., & Summers, R. M. (2021). Opportunistic screening at abdominal CT: Use of automated body composition

- biomarkers for added cardiometabolic value. *RadioGraphics*, 41(2), 524–542. <https://doi.org/10.1148/rg.2021200056>
- Pickhardt, P. J., Graffy, P. M., Zea, R., Lee, S. J., Liu, J., Sandfort, V., & Summers, R. M. (2020). Automated CT biomarkers for opportunistic prediction of future cardiovascular events and mortality in an asymptomatic screening population: A retrospective cohort study. *The Lancet Digital Health*, 2(4), e192–e200. [https://doi.org/10.1016/S2589-7500\(20\)30025-X](https://doi.org/10.1016/S2589-7500(20)30025-X)
- Ram, U. S., Pogue, J. A., Soike, M., Pfister, N. T., Jacob, R., & Cardenas, C. E. (2025). Assessing quantitative performance and expert review of multiple deep learning-based frameworks for computed tomography-based abdominal organ auto-segmentation. *Intelligent Oncology*, 1(2), 160–171. <https://doi.org/10.1016/j.intonc.2025.03.003>
- Sahiner, B., Pezeshk, A., Hadjiiski, L. M., Wang, X., Drukker, K., Cha, K. H., Summers, R. M., & Giger, M. L. (2019). Deep learning in medical imaging and radiation therapy. *Medical Physics*, 46(1), e1–e36. <https://doi.org/10.1002/mp.13264>
- Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R. R., & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, 10, 12598. <https://doi.org/10.1038/s41598-020-69250-1>
- Vasey, B., Nagendran, M., Campbell, B., Clifton, D. A., Collins, G. S., Denaxas, S., Denniston, A. K., Faes, L., Geerts, B., Ibrahim, M., Liu, X., Mateen, B. A., Mathur, P., McCradden, M. D., Morgan, L., Ordish, J., Rogers, C., Saria, S., Ting, D. S. W., ... McCulloch, P. (2022). Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ*, 377, e070904. <https://doi.org/10.1136/bmj-2022-070904>
- Wasserthal, J., Breit, H.-C., Meyer, M. T., Pradella, M., Hinck, D., Sauter, A. W., Heye, T., Boll, D. T., Cyriac, J., Yang, S., Bach, M., & Segeroth, M. (2023). TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. *Radiology: Artificial Intelligence*, 5(5), e230024. <https://doi.org/10.1148/ryai.230024>
- Yamashita, R., Mittendorf, A., Zhu, Z., Fowler, K. J., Santillan, C. S., Sirlin, C. B., Bashir, M. R., & Do, R. K. G. (2020). Deep convolutional neural network applied to the Liver Imaging Reporting and Data System (LI-RADS) version 2014 category classification: A pilot study. *Abdominal Radiology*, 45(1), 24–35. <https://doi.org/10.1007/s00261-019-02306-7>
- Yasaka, K., Akai, H., Abe, O., & Kiryu, S. (2018). Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: A preliminary study. *Radiology*, 286(3), 887–896. <https://doi.org/10.1148/radiol.2017170706>
- Zwanenburg, A., Vallières, M., Abdalah, M. A., Aerts, H. J. W. L., Andrearczyk, V., Apte, A., Ashrafinia, S., Bakas, S., Beukinga, R. J., Boellaard, R., Bogowicz, M., Boldrini, L., Buvat, I., Cook, G. J. R., Davatzikos, C., Depeursinge, A., Desseroit, M.-C., Dinapoli, N., Dinh, C. V., ... Löck, S. (2020). The Image Biomarker Standardization Initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*, 295(2), 328–338. <https://doi.org/10.1148/radiol.2020191145>

KAS-İSKELET SİSTEMİ RADYOLOJİSİNDE YAPAY ZEKA: KIRIK TESPİTİ VE MORFOMETRİK ÖLÇÜMLER

Elif Nisa Ünlü*

GİRİŞ

Kas-iskelet sistemi radyolojisi, günlük pratikte önemli bir yer tutan ve başta radyografi, manyetik rezonans görüntüleme (MRG) ve bilgisayarlı tomografi (BT) olmak üzere görüntüleme yöntemi çeşitliliği nedeniyle büyük miktarda veri üreten bir alt uzmanlık alanıdır. Bu yoğun veri akışı içerisinde tanısal doğruluğun korunması, raporlama süresinin yönetilmesi ve ölçümlerin standardizasyonu günümüzde giderek daha fazla önem kazanmaktadır. Bu nedenle yapay zeka (YZ), özellikle derin öğrenme yöntemlerinin gelişmesiyle kas-iskelet sistemi radyolojisinde önemli bir araştırma ve klinik uygulama alanı haline gelmiştir.

Kas-iskelet sistemi radyolojisinde YZ, tek bir teknoloji ya da tek bir klinik amaçtan oluşmaz. Güncel uygulamalar; kırık veya lezyon saptama, anatomik lokalizasyon, sınıflama, segmentasyon, nicel ölçüm, görüntü rekonstrüksiyonu ve iş listesi önceliklendirmesi gibi birbirinden farklı görevleri kapsar. Bu görevlerin her biri farklı bir referans standart ve performans ölçütü gerektirir. Örneğin görüntü düzeyinde kırık varlığını sınıflayan bir model ile kırık hattını işaretleyen bir model aynı klinik soruyu yanıtlamaz; benzer biçimde yüksek segmentasyon benzerliği, elde edilen açının veya kırık kalınlığının klinik olarak doğru olduğu anlamına tek başına gelmez (Guermazi vd., 2024; Park ve Han, 2018).

YZ'nin kas-iskelet görüntülemesindeki en yoğun kullanım alanı, radyografilerde kırık saptamaya yardımcı olmasıdır. İkinci ana alan olan morfometrik analiz ise görüntüdeki anatomik yapıların konum, şekil, uzunluk, açı, alan veya hacim gibi nicel özelliklerinin belirlenmesini amaçlar. Bununla birlikte, algoritmanın laboratuvar ortamındaki başarısı doğrudan klinik yarar anlamına gelmez. Olgu seçimi, prevalans, cihaz ve protokol farklılıkları, kullanıcı davranışı, veri kayması ve hatalı güven

*Prof. Dr., Düzce Üniversitesi, Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, nisaunlu@yahoo.com, ORCID: <https://orcid.org/0000-0002-4801-8854>

duygusu klinik performansı deęiřtirebilir. Bu nedenle YZ, özerk bir raporlama mekanizması olarak deęil, radyoloęun sorumluluęunu ve kalite kontrolünü koruyan bir klinik karar destek bileřeni olarak konumlandırılmalıdır (Recht vd., 2020; Omoumi vd., 2021).

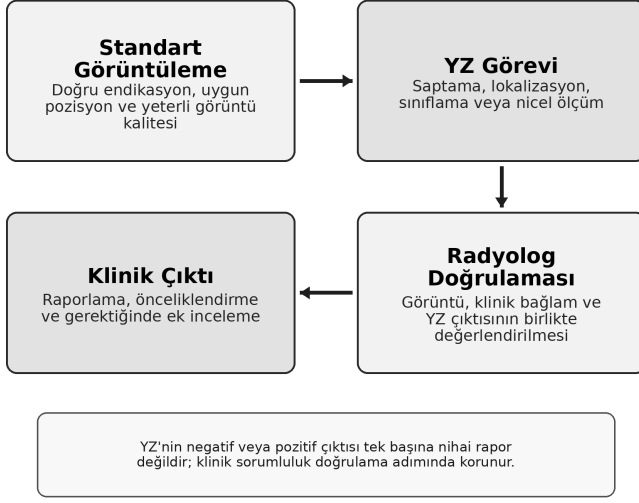
Bu bölümde otomatik kırık tespiti ile morfolometrik ölçümlerin kanıt düzeyi, ölçüm ilkeleri ve klinik sınırlılıkları ele alınmaktadır. MRG'de rekonstrüksiyon, sentetik BT ve yapısal lezyon sınıflaması, temel odaęı destekleyen iliřkili uygulamalar olarak daha sınırlı kapsamda tartiřılmıştır. YZ sistemlerinin klinik etkisi, yalnızca doęruluk ölçütleriyle deęil; raporlama süresi, geri çağırma veya ek inceleme oranı, kullanıcı hataları ve hasta bakımına etkisiyle birlikte deęerlendirilmelidir.

OTOMATİK KIRIK TESPİTİ

Klinik Görev ve İş Akışı

Kırık tespiti için geliştirilen bir sistem dört farklı rolden birini üstlenebilir: görüntüleri olası kırık açısından iş listesinde öne almak, inceleme sırasında eş zamanlı uyarı vermek, radyolog deęerlendirmesinden sonra ikinci okuyucu olarak çalışmak veya raporlama öncesinde kalite güvencesi sağlamak. Aynı algoritmanın performansı bu kullanım senaryolarında farklı sonuç doęurabilir. Önceliklendirme sisteminde duyarlılık ve gecikme süresi öne çıkarken, eş zamanlı karar desteęinde yanlış pozitif uyarıların kullanıcı yorgunluęuna etkisi de önemlidir. Negatif YZ çıktısı, özellikle klinik kuřkusu yüksek hastada görüntünün bağımsız deęerlendirilmesi gereęini ortadan kaldırmaz.

Algoritmanın ürettięi ısı haritası veya sınırlayıcı kutu, bulgunun anatomik lokalizasyonunu açıklamaya yardımcı olabilir; ancak bu görseller güvenilir bir belirsizlik ölçüsü deęildir. Benzer řekilde kırık sınıflaması, yer deęiřtirme, eklem içi uzanım ve eşlik eden çıkık gibi tedaviyi deęiřtiren ayrıntıların tümünü otomatik olarak kapsamayabilir. YZ'nin klinik deęeri, hangi görevi ne ölçüde yerine getirdięinin açık biçimde tanımlanmasına baęlıdır.



Şekil 1: Kas-iskelet sistemi görüntülemesinde YZ destekli klinik iş akışı.

Erişkin Radyografilerinde Kırık Saptama

Radyografide kırık saptama çalışmalarının toplu değerlendirmesinde, algoritmaların ve klinisyenlerin test kümelerindeki duyarlılık ve özgüllükleri genel olarak birbirine yakın bulunmuştur. Kuo ve arkadaşlarının sistematik derleme ve meta-analizinde 42 çalışma ve 55.061 görüntü değerlendirilmiş; iç testlerde YZ için havuzlanmış duyarlılık ve özgüllük sırasıyla yaklaşık %92 ve %91, klinisyenler için %91 ve %92 olarak bildirilmiştir. Dış test sonuçlarında da benzer bir düzey gözlenmekle birlikte çalışmaların yaklaşık yarısında yüksek yanlışlık riski saptanmıştır. Bu sonuç, YZ'nin kırık saptamada yararlı bir yardımcı olabileceğini destekler; ancak farklı merkez ve gerçek yaşam popülasyonlarında aynı performansın garanti edildiğini göstermez (Kuo vd., 2022).

Erken dönem çalışmalar, derin öğrenme modellerinin çok bölgesel travma radyografilerinde insan okuyuculara yaklaşabildiğini göstermiştir (Olczak vd., 2017). Daha sonraki okuyucu çalışmalarında YZ yardımı, özellikle gözden kaçma olasılığı bulunan kırıklarda duyarlılığı artırmış ve bazı iş akışlarında değerlendirme süresini kısaltmıştır. Örneğin kontrollü çok okuyuculu bir çalışmada YZ desteği okuyucu duyarlılığında 10,4, özgüllüğünde 5,0 yüzde puanlık artış ve inceleme süresinde ortalama 6,3 saniyelik azalma sağlamıştır. Bununla birlikte bu çalışmada olgu kümesi yapay olarak dengelenmiş, kırık prevalansı %50 olarak yapay biçimde belirlenmiş ve bu nedenle çalışma örnekleme gerçek klinik popülasyonu

temsil edememiştir (Guermazı vd., 2022). Bu nedenle sayısal kazanımlar klinik ortama doğrudan taşınmamalıdır.

YZ desteğinin yararı, okuyucunun deneyimine ve anatomik bölgeye göre değişebilir. Model, ince kortikal kırıklar, üst üste binen yapılar, osteopeni, alçı veya implant, ameliyat sonrası değişiklikler ve alışılmadık projeksiyonlarda hata yapabilir. Ayrıca eski kırık, aksesuar ossikül ve büyüme varyantları yanlış pozitif uyarıya neden olabilir. Kontrollü çalışmalar YZ'nin okuyucu performansını artırabileceğini gösterse de etkinin gerçek zamanlı klinik uygulamada sürdürülebilirliği; yerel prevalans, çekim kalitesi ve uyarıların iş istasyonunda nasıl sunulduğuyula birlikte ölçülmelidir (Lindsey vd., 2018; Canoni-Meynet vd., 2022).

Pediatric Kırıklar

Pediatric radyografiler, büyüme plakları, sekonder ossifikasyon merkezleri ve yaşa bağlı anatomi nedeniyle erişkin verisinden ayrı değerlendirilmelidir. Pediatric apendiküler iskelet kırıklarının otomatik saptanmasına ilişkin çalışmalar umut verici performans bildirmiştir. Ancak yaş dağılımı, anatomik bölge ve kırık tipi gibi alt gruplar arasında performans farklılıkları görülebilir ve özellikle daha az temsil edilen kırık tiplerinde model başarısının değerlendirilmesi güçleşebilir (Hayashi vd., 2022). Modelin yalnızca genel doğruluğunun verilmesi yeterli değildir; özellikle toddlers kırıkları, büyüme plağı yaralanmaları, dirsek travması ve çok küçük avulsiyonlar için alt grup performansı ayrıca değerlendirilmelidir. Bu nedenle çocuklarda negatif YZ çıktısının klinik muayene ve uygun takip görüntülemesinin yerini almaması özellikle önemlidir.

BT'de Vertebral Kırıklar ve Fırsatçı Değerlendirme

Toraks veya abdomen BT incelemelerinde vertebral kompresyon kırıkları, incelemenin birincil endikasyonu dışında kalabilir. Otomatik vertebra numaralandırma, korpus segmentasyonu ve yükseklik ölçümü bu bulguların sistematik biçimde taranmasına yardımcı olabilir. Bununla birlikte dejeneratif kama biçimi, Scheuermann hastalığı, eski travma ve kötü hizalanmış reformasyonlar yalancı sonuçlara yol açabilir. Bu nedenle algoritmanın vertebral düzey ataması ve işaretlediği son plaklar görsel olarak doğrulanmalıdır.

Fırsatçı BT değerlendirmesi, rutin klinik endikasyonlarla elde edilmiş görüntülerden ek nicel görüntüleme biyobelirteçlerinin elde edilmesine olanak sağlamaktadır. Bu kapsamda vertebral kemik atenüasyonu, kas

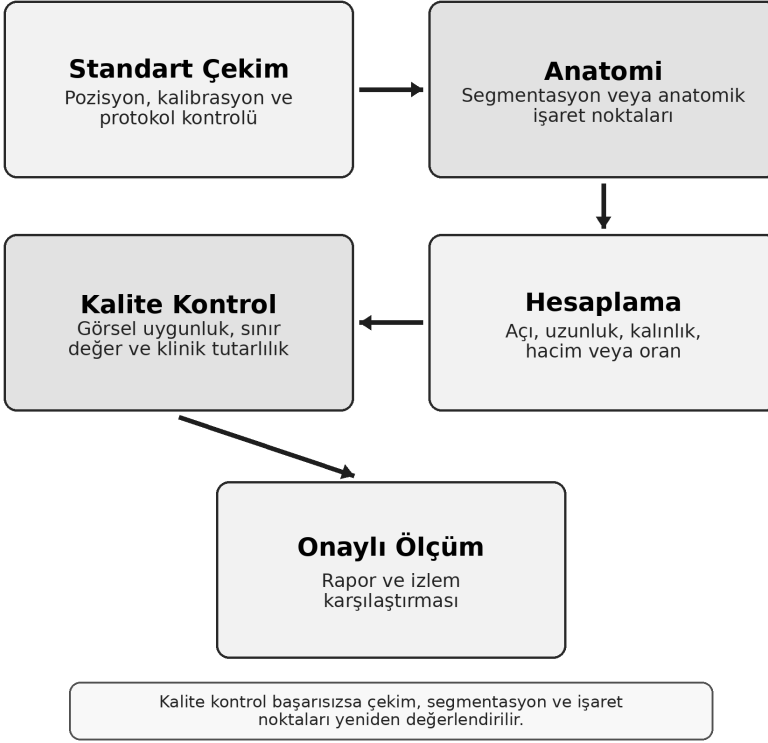
atenüasyonu ve yağ dağılımı gibi parametreler otomatik olarak ölçülebilmekte ve osteoporotik kırık riskinin değerlendirilmesinde kullanılabilir. Otomatik BT biyobelirteçlerinin gelecekteki majör osteoporotik kırık riskini öngörmede yararlı olabileceği bildirilmiştir (Pickhardt vd., 2020). Ancak bu tür bir çıktı osteoporoz tanısını tek başına koydurmaz ve çift enerjili X-ışını absorbsiyometrisi, klinik risk faktörleri veya uygun laboratuvar değerlendirmesinin yerine geçmez. Klinik uygulama için eşiklerin cihaz, kVp, kontrast fazı ve popülasyona göre doğrulanması; pozitif sonucun hangi yönlendirme sürecini başlatacağının önceden tanımlanması gerekir.

MORFOMETRİK ÖLÇÜMLER

Morfometrinin Teknik Temeli ve Performans Ölçütleri

Otomatik morfometri genellikle iki aşamalı bir süreç olarak ele alınabilir. İlk aşamada kemik, kırık veya ilgili yumuşak doku yapıları segmente edilir ya da anatomik işaret noktaları (landmark) belirlenir. İkinci aşamada ise bu yapılardan açı, uzunluk, alan, hacim, oran veya kalınlık gibi morfometrik parametreler hesaplanır. Nihai ölçüm hatası yalnızca yapay zekâ modelinin segmentasyon veya landmark belirleme performansından kaynaklanmayıp; görüntüleme ve çekim pozisyonu, cihaz kalibrasyonu, anatomik varyasyonlar ve kullanılan ölçüm tanımına bağlı olarak da değişebilir. Bu nedenle otomatik morfometri çalışmalarında yalnızca segmentasyon doğruluğunu gösteren Dice benzerlik katsayısının raporlanması yeterli değildir. Klinik ölçümün ortalama mutlak hatası (MAE), ölçüm yanlılığı (bias), sınıf içi korelasyon katsayısı (ICC) ve Bland–Altman analizi ile belirlenen uyum sınırları gibi ölçüm doğruluğu ve uyumunu değerlendiren parametrelerin de raporlanması gerekir.

Tekrarlanabilirlik ile doğruluk birbirinden ayrılmalıdır. Bir sistem aynı hatalı sınırı her seferinde seçerek yüksek tekrarlanabilirlik gösterebilir. Klinik kullanıma uygun bir morfometri sistemi, uzman referansına yakınlığın yanı sıra ardışık incelemelerde değişimi güvenilir biçimde ölçebilmelidir. Ölçüm öncesi görüntü kalitesi kontrolü, sonuç sonrası anatomik uygunluk kontrolü ve başarısız olguların kullanıcıya açıkça bildirilmesi temel güvenlik bileşenleridir.



Şekil 2: Otomatik morfometrik ölçümde kalite kontrol döngüsü.

Omurga ve Vertebral Morfometri

Omurga uygulamaları vertebral seviye belirleme, korpus yükseklikleri, kama oranı, sagittal denge ve skolyoz açılarının otomatik ölçümünü içerebilir. Özellikle Cobb açısının ölçümünde, vertebraların doğru şekilde belirlenmesi ve ilgili eğrinin tanımlanması ölçüm doğruluğunun temel bileşenleridir. Derin öğrenme tabanlı sistemlerin spinal radyograflardan Cobb açısını hızlı ve yüksek düzeyde uyum göstererek otomatik olarak ölçebildiği farklı çalışmalarda gösterilmiştir (Suri vd., 2023; Wong vd., 2023,). Ancak algoritmanın iki çizgi arasındaki açıyı doğru hesaplaması, yanlış son vertebra seçimini telafi etmez. Bu nedenle otomatik ölçüm; işaret noktaları ve kullanılan vertebralar görüntü üzerinde görünür olacak şekilde sunulmalı, gerektiğinde kullanıcı düzeltmesine izin vermeli ve düzeltme yapıldığında nihai raporda hangi değer kullanıldığı izlenebilmelidir.

Vertebral kırık morfometrisinde nicel yükseklik kaybı yararlı olmakla birlikte, görsel yarı nicel değerlendirme ile aynı kavram değildir. Dejeneratif değişiklikler, fizyolojik anterior kama, skolyoz ve kötü pozisyon yüksekliği

etkileyebilir. Otomatik oranlar, kırık lehine bir uyarı üretebilir; fakat kortikal kırık veya end plato kırığı gibi nitel bulgular ve önceki incelemeyle karşılaştırma radyolog tarafından değerlendirilmelidir.

Alt Ekstremitte Aksı ve Kalça Ölçümleri

Ayakta uzun bacak radyografilerinde kalça-diz-ayak bileği açısı ve femoral anatomik-mekanik açı gibi alt ekstremitte hizalanma parametreleri otomatik olarak belirlenen anatomik işaret noktalarından hesaplanabilir. Tam otomatik bir sistemin femur ve tibia segmentasyonunda yüksek benzerlik, açısal ölçümlerde yüksek uyum ve manuel ölçüme kıyasla daha kısa işlem süresi sağladığı gösterilmiştir (Schock vd., 2021). Buna karşın rotasyon, fleksiyon kontraktürü, pelvik obliklik, ekstremitte uzunluk farkı, görüntü birleştirme hatası ve ağırlık verme durumu ölçümün klinik anlamını değiştirebilir. YZ, standardize edilmemiş çekimin oluşturduğu biyomekanik hatayı ortadan kaldırmaz.

Kalça radyografilerinde lateral merkez-kenar açısı, Tönnis açısı, Sharp açısı ve femur başı örtünmesi gibi ölçümler otomatikleştirilebilir. Dış doğrulama çalışmaları, otomatik ölçümlerin hızlı gerçekleştirilebildiğini ve özellikle kalça displazisinin değerlendirilmesinde kullanılan temel radyolojik ölçümlerde uzman ölçümleriyle iyi ve mükemmel düzeyde uyum gösterebildiğini bildirmiştir (Archer vd., 2022). Bununla birlikte sınırda displazi, asferik femur başı, ameliyat sonrası anatomi ve yetersiz pelvik pozisyon olguları özel hata analizi gerektirir. Bir ölçümün istatistiksel olarak yüksek ICC göstermesi, klinik karar eşiğinin yakınındaki birkaç derecelik hatanın önemsiz olduğu anlamına gelmez. Bu nedenle otomatik ölçüm sistemlerinin değerlendirilmesinde yalnızca korelasyon veya uyum katsayılarının değil, mutlak ölçüm hatasının, sistematik yanlılığın ve klinik olarak kabul edilebilir hata sınırlarının da dikkate alınması önemlidir.

Kemik Yaşı

El-el bilek radyografisinden kemik yaşı tahmini, tekrarlayan ve iyi tanımlanmış bir görev olduğu için otomasyona elverişlidir. Derin öğrenme modelleri uzman değerlendirmesine yakın hata düzeyleri bildirmiş ve okuyucu iş akışını hızlandırma potansiyeli göstermiştir (Larson vd., 2018). Yine de kemik yaşı; biyolojik cinsiyet, kronolojik yaş, puberte evresi, etnik ve sosyoekonomik dağılım ile kullanılan atlas yönteminden etkilenir. Modelin eğitim popülasyonu, geçerli yaş aralığı ve belirsizliği raporda görünür olmalıdır. Endokrin bozukluklar, iskelet displazileri veya belirgin

asimetri gibi atlas dışı durumlarda otomatik sonucun uzman değerlendirmesi olmadan kullanılmaması gerekir.

MRG VE DİĞER NİCEL UYGULAMALAR

MR Rekonstrüksiyonu ve Sentetik BT

Derin öğrenme tabanlı süper çözünürlük ve rekonstrüksiyon yöntemleri, daha düşük çözünürlüklü ve kalın kesitli MR görüntülerinden yüksek çözünürlüklü görüntüler elde ederek çekim süresinin azaltılmasını amaçlamaktadır. Kas-iskelet MR çalışmalarında, süper çözünürlük yöntemleriyle yüksek çözünürlüklü görüntülere yakın görüntü kalitesi ve tanısal kalite düzeyinin korunabileceği gösterilmiş, ayrıca daha kısa çekim sürelerine yönelik potansiyel ortaya konmuştur (Chaudhari vd., 2018). Bununla birlikte piksel benzerliği veya okuyucu beğenisi, küçük kırıkdak defekti, ince tendon yırtığı ya da kırık hattının korunmasını tek başına garanti etmez. Rekonstrüksiyon ağı ince ayrıntıyı baskılayabilir veya görüntüde bulunmayan bir ayrıntı üretebilir. Bu nedenle hızlandırma katsayısı için evrensel bir güvenli eşik verilmemeli; her sekans, anatomik bölge ve klinik soru, tanısal ve klinik sonuçlarımlar üzerinden doğrulanmalıdır.

MRG'den sentetik BT üretimi, sakroiliak eklem gibi kemik yapının önemli olduğu alanlarda iyonizan radyasyon kullanmadan BT benzeri kemik kontrastı sağlamayı hedefler. Sakroiliit kuşkusunda erozyon, skleroz ve ankiloz gibi yapısal lezyonların saptanmasına ilişkin çalışmalar umut vericidir (Jans vd., 2021). Yine de sentetik BT gerçek BT'nin fiziksel atenüasyon ölçümünü üretmez; cihaz, sekans ve model bağımlıdır. Sonuçlar, doğrulandığı endikasyon dışında 'BT kalitesi' veya 'BT eşdeğeri' olarak genellenmemelidir.

Diz MRG'de Yapısal Lezyonlar

Diz MRG'de ön çapraz bağ ve menisküs yırtığı gibi bulguların sınıflanması, YZ araştırmalarının yoğunlaştığı alanlardandır. MRNet çalışması, çok düzlemli MR serilerinden genel anormallik, ön çapraz bağ yırtığı ve menisküs yırtığı sınıflaması için bir derin öğrenme modeli geliştirmiş ve geriye dönük bir çerçeve sunmuştur (Bien vd., 2018). Cerrahiye referans alan bir çalışmada menisküs yırtığı saptama performansı radyologlarla karşılaştırılmıştır (Fritz vd., 2020). Bununla birlikte sekans çeşitliliği, parsiyel yırtıklar, ameliyat sonrası değişiklikler, lezyonun yer ve morfolojisi ile eşlik eden bulgular model performansını etkiler. Sınıflama

çıktısı, sistematik bir MR değerlendirmesinin veya ayrıntılı radyoloji raporunun yerine geçmez.

Kıkırdak ve Osteoartrit

Osteoartrit değerlendirmesinde modalite ve ölçüt ayrımı korunmalıdır. Kellgren-Lawrence derecelendirmesi radyografiye dayanırken, kıkırdak kalınlığı ve hacmi gibi kantitatif ölçümler ile eklemdaki farklı yapısal değişiklikleri değerlendiren yarı nicel skorlamalar MR üzerinden gerçekleştirilebilir. Bu kavramların aynı algoritmik görev gibi sunulması yöntemsel karışıklık yaratabilir. T2 haritalarından osteoartrit sınıflaması ve otomatik kıkırdak segmentasyonu araştırma düzeyinde umut verici sonuçlar vermiştir (Pedoia vd., 2019; Wirth vd., 2021). Ancak yüksek Dice katsayısı, özellikle ince kıkırdak tabakasındaki sınır hatalarının kalınlık ve hacim hesabına etkisini gizleyebilir. Uzunlamasına çalışmalarda sistematik ofset, test-tekrar test değişkenliği ve gerçek biyolojik değişimin ayırt edilmesi gerekir.

KLİNİK DOĞRULAMA, İŞ AKIŞI VE HATA YÖNETİMİ

Uygun Performans Ölçütü ve Referans Standart

YZ araştırmasının değerlendirme yöntemi klinik görevle eşleşmelidir. Görüntü veya hasta düzeyinde sınıflama için duyarlılık, özgüllük, pozitif ve negatif öngörü değerleri ile alıcı işletim karakteristiği (ROC) eğrisi altında kalan alan kullanılabilir. Pozitif ve negatif öngörü değerlerinin hastalık prevalansına bağlı olduğu unutulmamalıdır. Birden fazla bulgunun lokalizasyonunda serbest yanıt ROC (FROC) yaklaşımı, segmentasyonda Dice katsayısına ek olarak HD95, nicel ölçümde ise ortalama mutlak hata, ICC ve Bland-Altman analizi daha uygundur. İş akışı çalışmalarında raporlama süresi, kritik bulgunun bildirim süresi, ek inceleme oranı ve kullanıcı hataları değerlendirilmelidir (Park ve Han, 2018).

Referans standart tek bir rapor olduğunda, rapor hataları model etiketine taşınabilir. Kırık çalışmalarında uzman uzlaşısı, ek projeksiyonlar, BT/MR, takip görüntüleme veya klinik kayıtlar gerektiğinde birleşik referans olarak kullanılmalıdır. Morfometride işaret noktası tanımı ve ölçüm protokolü önceden belirlenmeli; uzmanlar arası değişkenlik sunulmalıdır. Belirsiz olguların veri kümesinden çıkarılması performansı yapay olarak yükseltebilir ve gerçek yaşamdaki kullanım alanını daraltabilir.

Dış Doğrulama, Genellenebilirlik ve Veri Kayması

Aynı kurumdan ayrılan bir test kümesi dış doğrulama değildir. Klinik güven için farklı merkez, cihaz, üretici, protokol, yaş grubu ve hastalık spektrumunda zamansal veya coğrafi dış doğrulama gerekir. Veri kümesindeki kırık prevalansı, olgu zorluğu ve dışlama ölçütleri açıkça bildirilmelidir. Eğitim ve test kümelerinde aynı hastaya ait farklı incelemelerin bulunması veri sızıntısına yol açabilir. CLAIM 2024 güncellemesi; veri bölünmesi, referans standart, eksik veriler, hata analizi ve genellenebilirlik gibi unsurların şeffaf raporlanmasına yönelik bir çerçeve sunmaktadır (Tejani vd., 2024).

Model kullanıma girdikten sonra cihaz yazılımı, protokol, hasta popülasyonu veya sevk pratiğindeki değişiklikler performansı düşürebilir. Bu nedenle doğrulama bir defalık kabul testi değildir. Yerel performansın izlenmesi, kritik alt gruplar için eşiklerin gözden geçirilmesi ve yazılım güncellemelerinin yeniden değerlendirilmesi gerekir. Ticari ürün seçiminde yalnızca düzenleyici işaret veya üretici tarafından bildirilen AUC değeri yeterli değildir; hedef kullanım, yerel verideki performans, entegrasyon maliyeti, veri güvenliği ve hatada sorumluluk birlikte değerlendirilmelidir (Omoumi vd., 2021).

İnsan-YZ Etkileşimi ve Hata Türleri

YZ yardımı insan hatasını azaltabileceği gibi yeni hata biçimleri de oluşturabilir. Otomasyon yanlılığı, kullanıcının algoritma önerisini yeterince sorgulamadan kabul etmesidir; devre dışı bırakma yanlılığı ise sık yanlış uyarı nedeniyle gerçek uyarıların göz ardı edilmesidir. Bir başka risk, YZ'nin işaretlemediği bölgelerin daha az incelenmesidir. Kullanıcı arayüzü, modelin kapsamını ve başarısız olabileceği durumları görünür kılmalı; algoritma çıktısı görüntünün önüne geçmemelidir (Recht vd., 2020).

Hata analizi yalnızca yanlış pozitif ve yanlış negatif sayısı ile sınırlı olmamalıdır. Hatalar anatomik bölge, kırık tipi, hasta yaşı, implant, çekim kalitesi ve klinik önem açısından sınıflandırılmalıdır. Morfometrik sistemlerde başarısız segmentasyonun ölçüme dönüşmesini önleyen mantıksal sınırlar kullanılabilir; ancak sınırlar dışında kalmayan hataların da görsel olarak yakalanması gerekir. En güvenli yaklaşım, YZ'nin belirsizlik veya başarısızlık durumunda açık uyarı üretmesi ve ölçümün düzenlenebilir işaret noktalarıyla sunulmasıdır.

Yerel Uygulama ve İzlem

Yerel uygulama öncesinde hedef klinik sorun, kullanıcı grubu ve beklenen yarar tanımlanmalıdır. Ardından geçmiş yerel olgularda sessiz doğrulama, kullanıcı kabul testi ve sınırlı prospektif pilot yapılabilir. Raporlama iş istasyonu, radyoloji bilgi sistemi ve görüntü arşivleme sistemi entegrasyonu; sonuçların hangi anda gösterileceği ve yazılım çalışmadığında iş akışının nasıl sürdürüleceği belirlenmelidir. Model çıktısının rapora otomatik aktarılması durumunda ölçümün doğrulandığını gösteren bir kullanıcı adımı bulunmalıdır.

Uygulama sonrasında duyarlılık ve özgüllük kadar uyarı sayısı, kullanıcı kabul oranı, düzeltilen ölçüm oranı, raporlama süresi ve kritik gecikmeler izlenmelidir. Önemli hata örnekleri düzenli klinik toplantılarda gözden geçirilmeli; güncelleme veya yeniden eğitim gereksinimi belgelenmelidir. Bu yaklaşım, YZ'yi statik bir yazılım satın alımı olmaktan çıkarıp sürekli kalite güvencesi gerektiren bir tıbbi cihaz süreci olarak ele alır.

Tablo 1: Kas-iskelet sistemi görüntülemesinde YZ uygulamalarının klinik kanıt haritası

Görev ve çıktı	Uygun değerlendirme	Kanıt olgunluğu	Başlıca sınırlılık
Erişkin travma radyografisinde kırık saptama ve lokalizasyon	Duyarlılık, özgüllük, FROC; okuyucu performansı; işlem süresi	Çok merkezli okuyucu çalışmaları ve meta-analiz mevcut	Seçilmiş olgu dağılımı, yüksek kırık prevalansı ve cihaz/merkez farkı
Pediyatrik radyografide kırık saptama	Yaş grubuna göre duyarlılık ve özgüllük; olgu düzeyi hata analizi	Geriye dönük dış testler mevcut; klinik kanıt daha sınırlı	Büyüme plağı, ossifikasyon varyantları ve yaş dağılımı
BT'de vertebral kırık ve fırsatçı kemik değerlendirmesi	Hasta düzeyi duyarlılık; ölçüm tekrarlanabilirliği ; klinik yönlendirme	Umut verici geriye dönük çalışmalar	Tarama popülasyonu ve klinik risk araçlarıyla entegrasyon eksikliği
Omurga ve alt ekstremitte morфометrisi	Ortalama mutlak hata, ICC ve Bland-Altman uyumu	Tek merkezli ve bazı dış doğrulama çalışmaları	Pozisyon, rotasyon, implant ve anatomik işaret noktası hatası
Kalça ölçümleri ve kemik yaşı	ICC/uyum sınırları; uzmanlar arası karşılaştırma	Belirli görevlerde dış doğrulama mevcut	Demografi, görüntü protokolleri ve uç olgularda genellenebilirlik
Diz MR'sinde lezyon sınıflaması	Lezyon düzeyi duyarlılık/özgüllük; AUC; cerrahi veya uzman referansı	Geriye dönük çok görevli modeller; prospektif kanıt sınırlı	Sekans bağımlılığı, küçük lezyonlar ve etiket belirsizliği

Görev ve çıktı	Uygun değerlendirme	Kanıt olgunluğu	Başlıca sınırlılık
Kıkırdak segmentasyonu ve nicel MR	Dice/HD95; kalınlık-hacim hatası; uzunlamasına tekrarlanabilirlik	Araştırma ve klinik çalışma düzeyinde güçlü teknik kanıt	Yüksek Dice değeri klinik ölçüm doğruluğunu tek başına garanti etmez
MR rekonstrüksiyonu ve sentetik BT	Görüntü kalitesi yanında lezyon saptama ve tanısal eşdeğerlik	Erken klinik doğrulama	İnce bulguların baskılanması veya yapay ayrıntı oluşması riski

Not: Kanıt olgunluğu; yayın türü, dış doğrulama ve klinik iş akışına ilişkin verilerin birlikte değerlendirilmesine dayalı nitel bir sınıflamadır. FROC: Serbest yanıt alıcı işletim karakteristiği; ICC: Sınıf içi korelasyon katsayısı; HD95: Yüzde 95 Hausdorff uzaklığı. Kaynaklar: Kuo vd. (2022), Guermazi vd. (2022, 2024), Larson vd. (2018), Schock vd. (2021), Bien vd. (2018), Wirth vd. (2021) ve Jans vd. (2021).

SONUÇ VE ÖNERİLER

Kas-iskelet sistemi radyolojisinde YZ'nin en güçlü yakın dönem katkısı, iyi tanımlanmış görevlerde radyoloğun dikkatini ve ölçüm tutarlılığını desteklemesidir. Travma radyografilerinde kırık saptama için meta-analiz ve okuyucu çalışmaları klinik yarar olasılığını desteklemektedir; ancak olgu seçimi, yanlışlık riski ve gerçek yaşam prevalansı sonuçların yorumunu sınırlar. Morfometrik sistemler omurga, alt ekstremitte, kalça, kemik yaşı ve kıkırdak değerlendirmesinde zaman kazandırabilir; fakat çekim standardizasyonu ve görsel kalite kontrol olmadan güvenilir değildir.

Klinik uygulamada şu ilkeler benimsenmelidir: sistemin hedef görevi ve kapsamı açıkça tanımlanmalı, dış doğrulama ve yerel kabul testi yapılmalı, performans klinik göreve uygun ölçütlerle raporlanmalı, belirsiz veya başarısız çıktılar görünür olmalı, pozitif ve negatif sonuçlar radyolog tarafından doğrulanmalı, yazılım güncellemeleri ve veri kayması sürekli izlenmelidir. YZ'nin değeri, radyoloğun yerini almasıyla değil, hatayı azaltan, ölçümü standartlaştıran ve klinik kararı şeffaf biçimde destekleyen bir iş akışına güvenli olarak entegre edilmesiyle ortaya çıkar.

KAYNAKÇA

- Archer, H., Reine, S., Alshaikhsalama, A., Wells, J., Kohli, A., Vazquez, L., Hummer, A., DiFranco, M. D., Ljuhar, R., Xi, Y., & Chhabra, A. (2022). Artificial intelligence-generated hip radiological measurements are fast and adequate for reliable assessment of hip dysplasia: An external validation study. *Bone & Joint Open*, 3(11), 877-884. <https://doi.org/10.1302/2633-1462.311.BJO-2022-0125.R1>
- Bien, N., Rajpurkar, P., Ball, R. L., Irvin, J., Park, A., Jones, E., Bereket, M., Patel, B. N., Yeom, K. W., Shpanskaya, K., Halabi, S., Zucker, E., Fanton, G., Amanatullah, D. F., Beaulieu, C. F., Riley, G. M., Stewart, R. J., Blankenberg, F. G., Larson, D. B., Jones, R. H., Langlotz, C. P., Ng, A. Y., & Lungren, M. P. (2018). Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet. *PLOS Medicine*, 15(11), e1002699. <https://doi.org/10.1371/journal.pmed.1002699>
- Canoni-Meynet, L., Verdot, P., Danner, A., Calame, P., & Aubry, S. (2022). Added value of an artificial intelligence solution for fracture detection in the radiologist's daily trauma emergencies workflow. *Diagnostic and Interventional Imaging*, 103(12), 594-600. <https://doi.org/10.1016/j.diii.2022.06.004>
- Chaudhari, A. S., Fang, Z., Kogan, F., Wood, J., Stevens, K. J., Gibbons, E. K., Lee, J., Gold, G. E., & Hargreaves, B. A. (2018). Super-resolution musculoskeletal MRI using deep learning. *Magnetic Resonance in Medicine*, 80(5), 2139-2154. <https://doi.org/10.1002/mrm.27178>
- Fritz, B., Marbach, G., Civardi, F., Fucentese, S. F., & Pfirrmann, C. W. A. (2020). Deep convolutional neural network-based detection of meniscus tears: Comparison with radiologists and surgery as standard of reference. *Skeletal Radiology*, 49(8), 1207-1217. <https://doi.org/10.1007/s00256-020-03410-2>
- Guermazi, A., Omoumi, P., Tordjman, M., Fritz, J., Kijowski, R., Regnard, N. E., Carrino, J., Kahn, C. E., Jr., Knoll, F., Rueckert, D., Roemer, F. W., & Hayashi, D. (2024). How AI may transform musculoskeletal imaging. *Radiology*, 310(1), e230764. <https://doi.org/10.1148/radiol.230764>
- Guermazi, A., Tannoury, C., Kompel, A. J., Murakami, A. M., Ducarouge, A., Gillibert, A., Li, X., Tournier, A., Lahoud, Y., Jarraya, M., Lacave, E., Rahimi, H., Pourchot, A., Parisien, R. L., Merritt, A. C., Comeau, D., Regnard, N. E., & Hayashi, D. (2022). Improving radiographic fracture recognition performance and efficiency using artificial intelligence. *Radiology*, 302(3), 627-636. <https://doi.org/10.1148/radiol.210937>
- Hayashi, D., Kompel, A. J., Ventre, J., Ducarouge, A., Nguyen, T., Regnard, N. E., & Guermazi, A. (2022). Automated detection of acute appendicular skeletal fractures in pediatric patients using deep learning. *Skeletal Radiology*, 51(11), 2129-2139. <https://doi.org/10.1007/s00256-022-04070-0>
- Jans, L. B. O., Chen, M., Elewaut, D., Van den Bosch, F., Carron, P., Jacques, P., Wittoek, R., & Jaremko, J. L. (2021). MRI-based synthetic CT in the detection of structural lesions in patients with suspected sacroiliitis: Comparison with MRI. *Radiology*, 298(2), 343-349. <https://doi.org/10.1148/radiol.2020201537>
- Kuo, R. Y. L., Harrison, C., Curran, T. A., Jones, B., Freethy, A., Cussons, D., Stewart, M., Collins, G. S., & Furniss, D. (2022). Artificial intelligence in fracture detection: A systematic review and meta-analysis. *Radiology*, 304(1), 50-62. <https://doi.org/10.1148/radiol.211785>
- Larson, D. B., Chen, M. C., Lungren, M. P., Halabi, S. S., Stence, N. V., & Langlotz, C. P. (2018). Performance of a deep-learning neural network model in assessing skeletal maturity on pediatric hand radiographs. *Radiology*, 287(1), 313-322. <https://doi.org/10.1148/radiol.2017170236>
- Lindsey, R., Daluiski, A., Chopra, S., Lachapelle, A., Mozer, M. C., Sicular, S., Hanel, D., Gardner, M., Gupta, A., Hotchkiss, R., & Potter, H. (2018). Deep neural network improves fracture detection by clinicians. *Proceedings of the National Academy of Sciences*, 115(45), 11591-11596. <https://doi.org/10.1073/pnas.1806905115>
- Olczak, J., Fahlberg, N., Maki, A., Razavian, A. S., Jilert, A., Stark, A., Sköldenberg, O., & Gordon, M. (2017). Artificial intelligence for analyzing orthopedic trauma radiographs: Deep learning algorithms-are they on par with humans for diagnosing fractures? *Acta Orthopaedica*, 88(6), 581-586. <https://doi.org/10.1080/17453674.2017.1344459>

- Omoumi, P., Ducarouge, A., Tournier, A., Harvey, H., Kahn, C. E., Jr., Louvet-de Verchère, F., Pinto dos Santos, D., Kober, T., & Richiardi, J. (2021). To buy or not to buy-evaluating commercial AI solutions in radiology (the ECLAIR guidelines). *European Radiology*, 31(6), 3786-3796. <https://doi.org/10.1007/s00330-020-07684-x>
- Park, S. H., & Han, K. (2018). Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology*, 286(3), 800-809. <https://doi.org/10.1148/radiol.2017171920>
- Pedioia, V., Lee, J., Norman, B., Link, T. M., & Majumdar, S. (2019). Diagnosing osteoarthritis from T2 maps using deep learning: An analysis of the entire Osteoarthritis Initiative baseline cohort. *Osteoarthritis and Cartilage*, 27(7), 1002-1010. <https://doi.org/10.1016/j.joca.2019.02.800>
- Pickhardt, P. J., Graffy, P. M., Zea, R., Lee, S. J., Liu, J., & Sandfort, V. (2020). Automated abdominal CT imaging biomarkers for opportunistic prediction of future major osteoporotic fractures in asymptomatic adults. *Radiology*, 297(1), 64-72. <https://doi.org/10.1148/radiol.2020200466>
- Recht, M. P., Dewey, M., Dreyer, K., Langlotz, C., Niessen, W., Prainsack, B., & Smith, J. J. (2020). Integrating artificial intelligence into the clinical practice of radiology: Challenges and recommendations. *European Radiology*, 30(6), 3576-3584. <https://doi.org/10.1007/s00330-020-06672-5>
- Schock, J., Truhn, D., Abrar, D. B., Merhof, D., Conrad, S., Post, M., Mittelstrass, F., Kuhl, C., & Nebelung, S. (2021). Automated analysis of alignment in long-leg radiographs by using a fully automated support system based on artificial intelligence. *Radiology: Artificial Intelligence*, 3(2), e200198. <https://doi.org/10.1148/ryai.2020200198>
- Suri A, Tang S, Kargilis D, et al. Conquering the Cobb Angle: A Deep Learning Algorithm for Automated, Hardware-Invariant Measurement of Cobb Angle on Radiographs in Patients with Scoliosis. *Radiology: Artificial Intelligence*. 2023;5(4):e220158. <https://doi.org/10.1148/ryai.220158>
- Tejani, A. S., Klontzas, M. E., Gatti, A. A., Mongan, J. T., Moy, L., Park, S. H., & Kahn, C. E., Jr. (2024). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 update. *Radiology: Artificial Intelligence*, 6(4), e240300. <https://doi.org/10.1148/ryai.240300>
- Wirth, W., Eckstein, F., Kemnitz, J., Baumgartner, C. F., Konukoglu, E., Fuerst, D., & Chaudhari, A. S. (2021). Accuracy and longitudinal reproducibility of quantitative femorotibial cartilage measures derived from automated U-Net-based segmentation of two different MRI contrasts: Data from the Osteoarthritis Initiative healthy reference cohort. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 34(3), 337-354. <https://doi.org/10.1007/s10334-020-00889-7>
- Wong JC, Reformat MZ, Parent EC, Stampe KP, Southon Hryniuk SC, Lou EH. Validation of an artificial intelligence-based method to automate Cobb angle measurement on spinal radiographs of children with adolescent idiopathic scoliosis. *Eur J Phys Rehabil Med*. 2023;59(4):535-542. DOI:10.23736/S1973-9087.23.08091-7

ONKOLOJİK GÖRÜNTÜLEMEDE YAPAY ZEKÂ: RADYOMİK VE TEDAVİ YANITININ ÖNGÖRÜLMESİ

Ahmet Kerem İmrek*

GİRİŞ

Onkolojik görüntüleme, tümörün anatomik yayılımını ve zaman içindeki değişimini invaziv olmayan biçimde değerlendirebilmesi nedeniyle kanser bakımının temel bileşenlerinden biridir. Bilgisayarlı tomografi (BT), manyetik rezonans görüntüleme (MRG) ve pozitron emisyon tomografisi (PET), yalnızca lezyonun görünür boyutunu değil; şekil, yoğunluk, sinyal dağılımı, metabolik aktivite ve tümör çevresi hakkında da bilgi taşır. Radyomik ve yapay zekâ yöntemleri, bu görüntülerdeki nicel örüntüleri sistematik olarak çözümleyerek konvansiyonel görsel değerlendirmenin ötesinde ölçülebilir biyobelirteçler üretmeyi amaçlar (Lambin vd., 2012; Gillies vd., 2016).

Bu alandaki çalışmaların doğru yorumlanabilmesi için üç farklı klinik görevin birbirinden ayrılması gerekir. Tedavi yanıtının öngörülmesi, tedavi başlamadan önce veya erken tedavi döneminde elde edilen verilerle gelecekteki bir sonucun tahmin edilmesidir. Tedavi yanıtının değerlendirilmesi, belirli bir zaman noktasında hastanın yanıt durumunun RECIST 1.1, iRECIST ya da PERCIST gibi tanımlı ölçütlerle sınıflandırılmasıdır. RECIST otomasyonu ise hedef lezyon seçimi, lezyon eşleştirme, çap ölçümü ve yanıt sınıfı hesaplama gibi değerlendirme adımlarının yazılım desteğiyle yürütülmesidir. Otomasyon, geleceği öngören bir model olmak zorunda değildir; öngörü modeli de RECIST ölçümünü otomatikleştirmek zorunda değildir.

Radyomik, kişiselleştirilmiş onkolojide hasta seçimi, erken yanıtın belirlenmesi, tedavinin yoğunlaştırılması veya azaltılması ve gereksiz toksisitenin önlenmesi açısından önemli bir araştırma alanıdır. Bununla birlikte, yüksek ayırt edicilik bildiren tek merkezli bir modelin klinik açıdan yararlı, genellenebilir veya nedensel bir biyobelirteç olduğu kendiliğinden

*Dr. Öğretim Üyesi, Düzce Üniversitesi, Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, dr.ahmetimrek@gmail.com, ORCID: <https://orcid.org/0000-0003-2576-4683>

kabul edilemez. Veri üretiminden klinik karara kadar her aşamada standardizasyon, uygun doğrulama, kalibrasyon ve gerçek iş akışında etki değerlendirmesi gerekir (Lambin vd., 2017; Bera vd., 2022).

Bu bölümde radyomik iş akışı, tedavi yanıt ölçütleri, yapay zekâ destekli RECIST uygulamaları, tedavi yanıtı öngörüsünde başlıca klinik kullanım alanları ve model geliştirme ilkeleri ele alınmaktadır. Amaç, yöntemin potansiyelini klinik kanıtın sınırlarıyla birlikte değerlendirmek ve araştırmadan rutin kullanıma geçiş için gerekli asgari koşulları tanımlamaktır.

RADYOMİK VE YAPAY ZEKÂ TEMELLERİ

Radyomik, Derin Radyomik ve Delta-Radyomik

Radyomik, tıbbi görüntülerde tanımlanan bir ilgi hacminden çok sayıda nicel özelliğin çıkarılması ve bu özelliklerin klinik ya da biyolojik sonuçlarla ilişkilendirilmesi sürecidir. El yapımı radyomik özellikler; birinci derece yoğunluk istatistikleri, şekil ve hacim ölçütleri ile gri düzey eş-oluşum, koşu uzunluğu ve benzeri matrislerden türetilen doku özelliklerini kapsar. Filtrelenmiş görüntülerden çıkarılan dalgacık ve Laplasyen-Gauss özellikleri farklı uzamsal ölçeklerdeki örüntüleri temsil edebilir. Bu değişkenler görüntüdeki heterojenliği nicelleştirir; ancak tek bir özelliğe doğrudan biyolojik anlam yüklemek çoğu zaman mümkün değildir (Liu vd., 2019; Scapicchio vd., 2021).

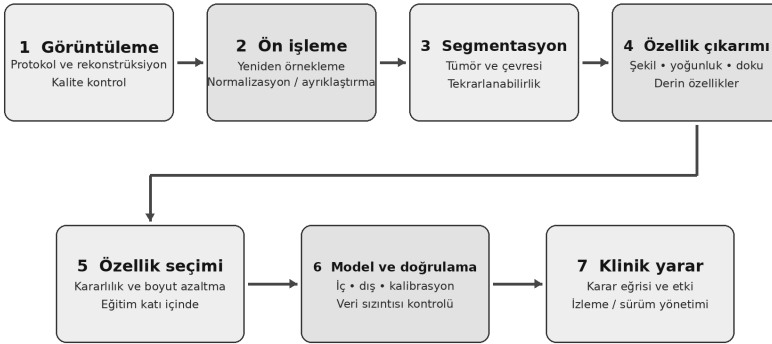
Derin radyomikte özellikler önceden tanımlanmış matematiksel formüllerle değil, çoğunlukla evrimsel sınır ağlarının ara katmanlarından öğrenilir. Uçtan uca yaklaşımlar segmentasyon, özellik öğrenme ve sınıflandırmayı aynı model içinde birleştirebilir. Bunun bedeli daha büyük örneklem gereksinimi, daha düşük açıklanabilirlik ve cihaz ya da protokol değişimine karşı duyarlılıktır. El yapımı ve derin özelliklerin birlikte kullanıldığı hibrit modeller teorik olarak tamamlayıcı bilgi sağlayabilir; ancak değişken sayısını artırdıkları için aşırı uyum ve veri sızıntısı riskini de büyütür.

Delta-radyomik, aynı hastanın farklı zaman noktalarındaki görüntülerinden çıkarılan özelliklerin mutlak ya da göreceli değişimini inceler. Amaç, tümörün tedaviye erken biyolojik yanıtını yalnızca çap değişiminden önce yakalayabilmektir. Bu yaklaşım immünoterapi, kemoradyoterapi ve bölgesel tedavilerde umut verici sonuçlar üretmiştir; buna karşın zamanlama, görüntü eşleştirme, rekonstrüksiyon ve segmentasyon farklılıkları biyolojik

değişim ile ölçüm gürültüsünü birbirine karıştırabilir. Güncel sistematik derlemeler, delta-radyomik çalışmalarının çoğunda yöntemsel titizlik ve bağımsız doğrulama düzeyinin sınırlı olduğunu göstermektedir (Nardone vd., 2024).

Radyomik İş Akışı

Güvenilir bir radyomik çalışma, klinik sorunun ve tahmin zamanının açıkça tanımlanmasıyla başlar. Görüntü elde etme ve rekonstrüksiyon parametreleri kaydedilir; ardından gerekirse voksel boyutu eşleştirme, yoğunluk normalizasyonu, gürültü azaltma ve gri düzey ayrıklaştırma uygulanır. Lezyon veya tümör çevresi elle, yarı otomatik ya da otomatik olarak bölütlenir. Özellik çıkarımından sonra kararsız ve birbiriyle yüksek korelasyonlu değişkenler elenir; özellik seçimi ile model eğitimi yalnızca eğitim verisi içinde yürütülür. Son aşama, iç ve dış doğrulama ile klinik yararın değerlendirilmesidir (van Griethuysen vd., 2017; Zwanenburg vd., 2020).



Şekil 1: Onkolojik radyomik iş akışının veri üretiminden klinik doğrulamaya kadar temel bileşenleri (yazar tarafından hazırlanmıştır).

Standardizasyon ve Tekrarlanabilirlik

Radyomik özellikler hasta biyolojisinin yanı sıra kesit kalınlığı, tüp akımı, rekonstrüksiyon çekirdeği, MRG sekansı, PET rekonstrüksiyonu, kontrast fazı ve hareket artefaktından etkilenebilir. Aynı lezyon farklı protokollerle görüntülendiğinde özellik dağılımı değişebilir. Bu nedenle yalnızca modelin sonuç ölçütlerini bildirmek yeterli değildir; görüntüleme parametreleri, ön işleme kararları ve özellik tanımları yeniden üretilebilir biçimde raporlanmalıdır. Tekrar çekim ve yeniden bölütleme analizleri,

hangi özelliklerin ölçüm değişkenliğine dirençli olduğunu göstermede yararlıdır (Traverso vd., 2018).

Segmentasyon, iş akışının en kritik belirsizlik kaynaklarından biridir. Sınırları düzensiz, nekrotik veya infiltratif lezyonlarda gözlemciler arası fark; şekil, yüzey ve doku özelliklerini önemli ölçüde değiştirebilir. Segmentasyon algoritmasının başarısı yalnızca örtüşme katsayısıyla değil, aşağı akım radyomik özellikler ve klinik model üzerindeki etkisiyle de incelenmelidir. Tümör çevresinin modele eklenmesi biyolojik olarak anlamlı olabilir; fakat çevresel halkanın kalınlığı gibi kararlar önceden tanımlanmadığında araştırmacı serbestlik derecesi artar.

Tekrarlanabilirlik ve yeniden üretilebilirlik aynı kavram değildir. Tekrarlanabilirlik, aynı koşullarda yinelenen çekim veya ölçümde benzer sonuç elde edilmesini; yeniden üretilebilirlik ise farklı cihaz, protokol, yazılım ya da gözlemcide sonucun korunmasını ifade eder. Fantom, tekrar çekim ve çoklu segmentasyon çalışmaları bu iki boyutu ayrı ayrı sınavabilir. Kararsız özellikleri yalnızca tüm kohortta hesaplanan bir eşik ile elemek de doğrulama verisinden bilgi sızdırabileceğinden, kararlılık temelli seçim eğitim verisi içinde yapılmalıdır. Özellik kararlılığı model kararlılığını garanti etmez; nihai tahminin farklı teknik koşullardaki değişimi ayrıca incelenmelidir.

Image Biomarker Standardisation Initiative, özellik tanımlarını ve hesaplama ayrıntılarını uyumlaştırmak için ortak bir çerçeve sunmuştur. PyRadiomics gibi açık kaynaklı yazılımlar uygulama saydamlığını artırırken, kullanılan sürüm ve parametre dosyasının paylaşılması yine de gereklidir (van Griethuysen vd., 2017; Zwanenburg vd., 2020). CLEAR kontrol listesi radyomik araştırmaların raporlanmasında klinik bağlam, veri, segmentasyon, özellik çıkarımı, modelleme ve açık bilim bileşenlerinin sistematik biçimde sunulmasını destekler (Kocak vd., 2023).

Çok merkezli verilerde ComBat benzeri uyumlaştırma yöntemleri merkez veya cihaz kaynaklı dağılım farklarını azaltabilir. Ancak uyumlaştırma, yetersiz protokol denetiminin yerine geçmez; biyolojik sinyali de zayıflatır ve sonuç değişkeniyle ilişkili merkez etkilerini maskeleyebilir. Parametrelerin tüm veri üzerinde öğrenilmesi test bilgisinin eğitim sürecine sızmasına yol açar. Bu nedenle uyumlaştırma yalnızca eğitim bölümünde öğrenilmeli, doğrulama verisine önceden tanımlı biçimde uygulanmalı ve uyumlaştırma öncesi-sonrası sonuçlar duyarlılık analiziyle gösterilmelidir (Orlhac vd., 2022).

TEDAVİ YANITININ GÖRÜNTÜLEMEYLE DEĞERLENDİRİLMESİ

RECIST 1.1

Solid tümörlerde yanıt değerlendirmesinin tarihsel temeli, ölçülebilir hastalığın boyut değişimine dayanır. İlk RECIST kılavuzu 2000 yılında yayımlanmış, RECIST 1.1 ile hedef lezyon sayısı, lenf nodu ölçümü ve progresyon tanımları güncellenmiştir (Therasse vd., 2000; Eisenhauer vd., 2009). RECIST 1.1'de başlangıçta en fazla beş hedef lezyon ve organ başına en fazla iki lezyon seçilir. Nodal olmayan hedef lezyonlarda uzun eksen, patolojik lenf nodlarında kısa eksen kullanılır; hedef dışı hastalık ve yeni lezyonlar ayrıca değerlendirilir.

Kısmi yanıt, hedef lezyon çapları toplamında başlangıca göre en az %30 azalma; progresif hastalık ise izlemdeki en düşük toplam değere göre en az %20 ve mutlak olarak en az 5 mm artış olarak tanımlanır. Yeni lezyon da progresyon kabul edilir. Bu eşiklerin doğru uygulanması için aynı lezyonların zaman içinde güvenilir biçimde eşleştirilmesi, ölçüm düzleminin tutarlı seçilmesi ve belirsiz yeni bulguların klinik bağlamla yorumlanması gerekir (Eisenhauer vd., 2009).

RECIST 1.1 basit, yaygın ve klinik araştırmalar arasında karşılaştırılabilir bir çerçeve sunar; buna karşın tümör heterojenliğini, nekroz veya kaviteasyon gibi iç değişiklikleri ve metabolik yanıtı doğrudan ölçmez. Hedef lezyon seçimi ile sınır belirleme gözlemci değişkenliğine açıktır. Ayrıca ölçülen birkaç lezyon, yaygın metastatik hastalıktaki toplam tümör yükünü her zaman temsil etmeyebilir. Bu sınırlılıklar, hacimsel ölçüm ve yapay zekâ desteğine yönelik araştırmaların temel gereksesidir.

Hacimsel yanıt değerlendirmesi, tek boyutlu çapın kaçırabileceği üç boyutlu değişimi daha duyarlı gösterebilir; ancak hacimdeki küçük farkların biyolojik yanıt mı yoksa segmentasyon değişkenliği mi olduğu her lezyon için aynı değildir. Hacim eşiği RECIST 1.1'in yerleşik çap eşiklerinden doğrudan türetilemez ve tümör şekline göre farklı sonuç verebilir. Bu nedenle hacimsel ölçümler araştırmalarda tamamlayıcı sonlanım olarak sunulmalı; kullanılan bölütleme yöntemi, en küçük güvenilir değişim ve başarısız otomasyon oranı açıkça raporlanmalıdır.

iRECIST ve PERCIST

İmmün kontrol noktası inhibitörleri, tümör yükünde geçici artış veya yeni lezyon sonrası yanıt gibi alışılmadık örüntüler oluşturabilir. iRECIST,

RECIST 1.1 ölçümlerini korurken ilk progresyonu doğrulanmamış progresif hastalık olarak sınıflandırır ve klinik açıdan stabil hastada sonraki değerlendirmeye doğrulama olanağı tanır. iRECIST öncelikle immünoterapi araştırmalarında veri standardizasyonu için geliştirilmiştir; klinik tedavi kararını tek başına belirleyen bir kural olarak yorumlanmamalıdır (Seymour vd., 2017).

PERCIST, 18F-FDG PET/BT'de metabolik aktivite değişimini standartlaştırmayı amaçlar ve yağsız vücut kitlesine göre düzeltilmiş tepe standartlaştırılmış tutulum değerini temel alır. Metabolik değişim, anatomik küçülmeden önce ortaya çıkabileceği için bazı tümör ve tedavilerde erken yanıt hakkında ek bilgi sağlayabilir. Bununla birlikte kan glukozu, enjeksiyon-görüntüleme aralığı, rekonstrüksiyon ve inflamasyon gibi etkenler ölçümü etkiler; tüm tümörler de yeterli FDG tutulumu göstermez (Wahl vd., 2009).

Tablo 1: Onkolojik tedavi yanıtında başlıca görüntüleme ölçütlerinin karşılaştırılması

Ölçüt	Temel ölçüm	Başlıca kullanım	Yanıt sınıfları	Temel sınırlılık
RECIST 1.1	BT/MRG'de hedef lezyon çapları; nodlarda kısa eksen	Solid tümörlerde anatomik yanıt	Tam, kısmi, stabil, progresif	Sınırlı lezyon örnekleme; heterojenlik ve metabolizma doğrudan ölçülmez
iRECIST	RECIST 1.1 ölçümleri ve progresyonun doğrulanması	İmmünoterapi araştırmaları	iCR, iPR, iSD, iUPD, iCPD	Klinik stabilite ve doğrulama gerekir; tedavi kararı için tek başına yeterli değildir
PERCIST	18F-FDG PET/BT'de SULpeak değişimi	Metabolik tedavi yanıtı	Tam, kısmi, stabil, progresif metabolik yanıt	Teknik standardizasyona ve FDG aviditesine bağımlıdır; inflamasyonla karışabilir

Not. RECIST 1.1 için Eisenhauer vd. (2009), iRECIST için Seymour vd. (2017) ve PERCIST için Wahl vd. (2009) esas alınmıştır.

Yapay Zekâ Destekli RECIST Otomasyonu

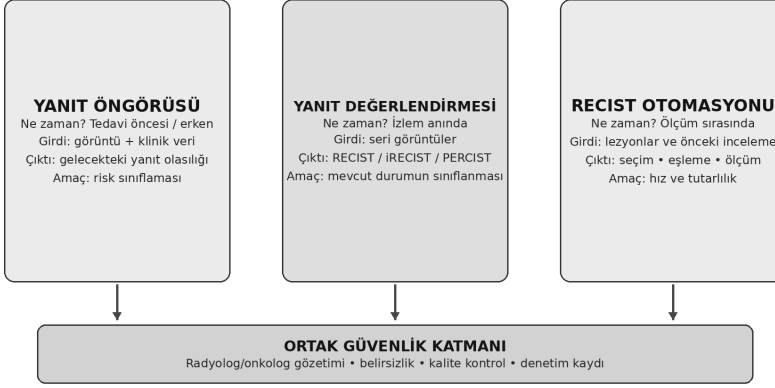
RECIST otomasyonu tek bir algorithmadan değil, ardışık görevlerden oluşur: aday lezyonların saptanması, ölçülebilirliğin belirlenmesi, hedef lezyon seçimi, önceki incelemeyle eşleştirme, çap veya hacim ölçümü, yeni lezyon saptama ve yanıt sınıfının hesaplanması. Bu adımlardan birindeki hata sonraki aşamalara taşınır. Özellikle birleşen lezyonlar, atelektaziye komşu akciğer tümörleri, tedavi sonrası skar ve sınırları belirsiz karaciğer metastazları klinik muhakeme gerektirir.

Çok merkezli ve çok gözlemcili bir BT çalışmasında otomatik uzun çap ve hacim ölçümlerinin izlem değerlendirmesini destekleyebileceği gösterilmiş; ancak performansın lezyon tipine ve insan düzeltmesine bağlı olduğu vurgulanmıştır (Dahm vd., 2024). Benzer biçimde, bir yapay zekâ sisteminin akciğer tümör yanıtı değerlendirmesinde ölçüm ve iş akışı verimliliğine katkı sağlayabileceği bildirilmiştir (Nie vd., 2024). Bu bulgular otomatik ölçümün uygulanabilirliğini gösterir; hedef lezyon seçimi ve klinik yanıt yorumunun tamamen çözüldüğünü göstermez.

Derin öğrenmeyle yarı otomatik tek yönlü akciğer tümörü ölçümü, radyolog ölçümleriyle uyumlu sonuçlar üretebilmiştir (Woo vd., 2021). Difüzyon ağırlıklı MRG'de karaciğer metastazlarının otomatik bölütlenmesine dayalı RECIST değerlendirmesi de araştırılmıştır; küçük, birleşen veya düşük kontrastlı lezyonlarda sınırlılıklar devam etmiştir (Liu vd., 2022). Hedef lezyon seçimini modelleyen çalışmalar ise radyologlar arasında başlangıçta bulunan değişkenliği yeniden üretme tehlikesini ortaya koymaktadır (Tareco Bucho vd., 2024).

Yapılandırılmamış rapor metninden RECIST yanıtını tahmin eden modeller, büyük retrospektif veri tabanlarında sonuç çıkarımını hızlandırabilir. Akciğer kanserinde PD-1 blokajı alan hastaların raporlarıyla geliştirilen bir derin öğrenme yaklaşımı bu kullanımın uygulanabilirliğini göstermiştir (Arbour vd., 2021). Ancak rapordan çıkarılan yanıt, görüntülerin yeniden merkezi değerlendirilmesinin eşdeğeri değildir; raporlama alışkanlıklarını ve mevcut ölçüm hatalarını da öğrenebilir. Araştırma veri seti oluşturma, kalite kontrol ve klinik karar desteği kullanım senaryoları bu nedenle ayrı doğrulama gerektirir.

Güvenli otomasyon, modelin ölçüm vermediği veya insan incelemesi istediği durumları da performans hesabına katmalıdır. Yalnızca algoritmanın başarıyla işlediği kolay olgular değerlendirilirse seçici başarısızlık gizlenir. Lezyon başına doğruluk, hasta başına nihai yanıt uyumu, gerekli düzeltme süresi ve klinik olarak önemli sınıf değişiklikleri birlikte raporlanmalıdır. Belirsizlik yüksek olduğunda sistemin otomatik sonucu durdurması, güvenilirlik eşiği ve kullanıcıya sunulan gerekçe önceden belirlenmelidir.



Şekil 2: Yanıt öngörüsü, yanıt değerlendirmesi ve RECIST otomasyonunun klinik görev bakımından ayrımı (yazar tarafından hazırlanmıştır).

TEDAVİ YANITININ ÖNGÖRÜLMESİNDE RADYOMİK

Tedavi Öncesi Öngörü

Tedavi öncesi radyomik modeller, başlangıç görüntüsündeki fenotipi gelecekteki patolojik, anatomik veya metabolik yanıtla ilişkilendirir. Örneğin heterojenlik, sınır düzensizliği ve tümör çevresi özellikleri hipoksi, nekroz, stromal yanıt veya immün mikroçevreyle ilişkili örüntüler taşıyabilir. Bununla birlikte görüntü özelliği ile biyolojik mekanizma arasındaki ilişki çoğu zaman dolaylıdır. Bir modelin yanıt verenlerle vermeyenleri ayırması, özelliğin nedensel bir belirteç olduğu veya tedavi seçimini değiştireceği anlamına gelmez (Bera vd., 2022).

Öngörü hedefi açık biçimde tanımlanmalıdır. Patolojik tam yanıt, RECIST yanıtı, progresyonsuz sağkalım ve genel sağkalım aynı sonlanım değildir; farklı biyolojik ve klinik süreçleri temsil eder. Görüntüleme ile sonuç arasındaki zaman aralığı, ara tedaviler ve sonuç değerlendirme yöntemi önceden belirtilmelidir. Tedavi seçimi hekim kararıyla yapılıyorsa, model tedavi etkisini değil hasta seçimine eşlik eden karıştırıcı değişkenleri öğrenebilir.

Başlangıç görüntüsünden geliştirilen bir imzanın en gerçekçi ilk kullanım alanı, tek başına tedavi belirlemekten çok risk sınıflaması ve araştırma zenginleştirmesidir. Yüksek riskli olduğu öngörülen hastalarda daha erken görüntüleme, ek biyobelirteç veya multidisipliner değerlendirme planlanabilir. Tedavi değişikliği öneren bir model için ise yalnızca ayırım

performansı değil, alternatif stratejiye göre net klinik yarar ve hasta sonuçlarında iyileşme gösterilmelidir.

Delta-Radyomik ve Erken Yanıt

Delta-radyomik, tedavi öncesi ve erken tedavi görüntüleri arasındaki özellik değişimini kullanır. Bu yaklaşım, başlangıç fenotipine ek olarak tümörün uygulanan tedaviye verdiği dinamik tepkiyi içerdiği için teorik olarak daha kişiseldir. Hücresellik, perfüzyon veya metabolizmadaki değişim çap azalmasından önce görülebilir. İmmünoterapi yanıtı üzerine yapılan derlemelerde delta özelliklerin umut verici olduğu, fakat küçük örneklem, heterojen zaman noktaları ve sınırlı dış doğrulama nedeniyle klinik kanıtın erken aşamada kaldığı bildirilmiştir (Abbas vd., 2023; Nardone vd., 2024).

Delta ölçümünün güvenilirliği iki zaman noktasının karşılaştırılabilirliğine bağlıdır. Aynı modalite, benzer protokol, uygun görüntü kaydı ve tutarlı segmentasyon olmadan ölçülen değişim teknik değişkenliği yansıtabilir. Yüzde değişim, başlangıç değeri sıfıra yakın özelliklerde kararsız olabilir; mutlak fark ise ölçekten etkilenir. Kullanılan delta tanımı, zaman penceresi ve eksik izlem görüntülerinin yönetimi önceden belirlenmelidir.

Erken yanıt modelinde ölümsüz zaman ve seçim yanlılığı da göz önünde bulundurulmalıdır. İkinci görüntüye ulaşamayan, tedaviyi erken bırakan veya hızla ilerleyen hastaların dışlanması modelin gerçek dünyadaki performansını olduğundan iyi gösterebilir. Bu nedenle başlangıçtaki uygun hasta kohortu, takip kayıpları ve görüntüleme zamanlamasından sapmalar akış şemasıyla raporlanmalıdır.

Çok Modlu Veri Bütünleştirilmesi

Radyomik özelliklerin yaş, evre, performans durumu, laboratuvar bulguları, histopatoloji ve moleküler verilerle birleştirilmesi, görüntünün tek başına yakalayamadığı bilgiyi modele ekleyebilir. Erken birleştirme tüm değişkenleri tek tabloda toplarken, geç birleştirme ayrı modellerin çıktılarını bütünleştirir. Derin öğrenme tabanlı ortak temsil yöntemleri daha karmaşık ilişkileri öğrenebilir; ancak örneklem gereksinimi ve açıklama gücü artar.

Metastatik hastalıkta birden fazla lezyonun nasıl birleştirildiği de model tanımının parçasıdır. En büyük lezyonun kullanılması basittir, fakat lezyonlar arası heterojenliği göz ardı eder. Ortalama veya maksimum özellik, toplam hacimle ağırlıklandırma ve lezyon temsillerini öğrenen çoklu örnek yaklaşımı farklı biyolojik varsayımlar taşır. Hastanın toplam tümör yükü,

organ dağılımı ve yeni lezyon gelişimi yalnızca tek bir hedef lezyonun fenotipinden çıkarılamaz. Birleştirme kuralı sonuca bakılarak seçilmemeli; hasta düzeyinde doğrulama yapılmalı ve lezyon sayısına bağlı performans değişimi gösterilmelidir.

Birleşik modelin yalnızca görüntü modelinden daha yüksek eğri altı alan vermesi yeterli değildir. Artışın dış doğrulamada sürmesi, kalibrasyonu bozmaması ve klinik karar eşiğinde net yarar sağlaması gerekir. Eksik moleküler testler veya düzensiz laboratuvar ölçümleri gerçek yaşamda sık olduğundan, eksik veri mekanizması ve imputasyon yöntemi eğitim-doğrulama ayrımı korunarak ele alınmalıdır. Modelin yalnızca tüm veri bileşenleri mevcut hastalarda çalışması erişilebilirliği azaltabilir.

Seçilmiş Klinik Uygulamalar

Küçük hücreli dışı akciğer kanserinde BT ve PET tabanlı radyomik, immün kontrol noktası inhibitörlerine yanıt ve sağkalımı öngörmek amacıyla yoğun biçimde araştırılmıştır. Sistemik derlemeler ayırt edici performansın umut verici olduğunu, ancak çalışmaların çoğunun retrospektif, tek merkezli ve yöntemsel kalite bakımından değişken olduğunu göstermektedir (Chen vd., 2021). Pseudoprogresyon, tedavi çizgisi, PD-L1 durumu ve eş zamanlı tedaviler hedef tanımını etkilediği için kohortlar arasında doğrudan karşılaştırma güçtür.

Lokal ileri rektum kanserinde temel klinik hedeflerden biri neoadjuvan kemoradyoterapi sonrası patolojik tam yanıtın öngörülmesidir. MRG radyomik çalışmaları T2 ağırlıklı görüntüler, difüzyon ve görünür difüzyon katsayısı haritalarını kullanmıştır. Geniş bir güncel meta-analiz yüksek özet performans bildirse de hasta seçimi, MRG alan gücü, sekanslar, bölütleme ve özellik çıkarımındaki heterojenliğin önemli olduğunu göstermiştir (Liao vd., 2025). Organ koruma kararları için bu modellerin prospektif, çok merkezli ve kalibre edilmiş biçimde doğrulanması gerekir.

Meme kanserinde dinamik kontrastlı MRG, difüzyon MRG ve PET tabanlı modeller neoadjuvan kemoterapiye patolojik yanıtı tahmin etmek için kullanılmaktadır. MRG ve makine öğrenmesi çalışmalarını birleştiren meta-analizler orta-yüksek ayırt edicilik bildirmiş, klinik ve patolojik değişkenlerin eklenmesinin yararlı olabileceğini düşündürmüştür (Liang vd., 2022). Moleküler alt tip, tedavi rejimi ve patolojik tam yanıt tanımındaki farklılıklar model aktarılabilirliğini sınırlayan temel etkenlerdir.

Hepatoselüler karsinomda BT ve MRG radyomik, transarteriyel kemoembolizasyon yanıtını tedavi öncesinde öngörmek için değerlendirilmiştir. Meta-analitik bulgular potansiyel bulunduğunu, buna karşın karaciğer fonksiyonu, tümör evresi, embolizasyon tekniği ve yanıt ölçütünün çalışmalar arasında değiştiğini göstermektedir (Feng vd., 2023). Model, yalnızca görüntü fenotipini değil, tedavi tekniği ve hasta seçimine ilişkin merkez uygulamalarını da öğrenebilir.

Baş-boyun kanserinde radyomik; kemoradyoterapi yanıtı, lokorejyonel kontrol ve toksisite öngörüsünde araştırılmaktadır. Primer tümör, nodal hastalık ve tedavi planlama görüntülerinden farklı bilgiler elde edilebilir. Bununla birlikte dış artefaktı, küçük ve düzensiz hacimler, HPV durumunun dağılımı ve çoklu sonlanımlar yöntemsel güçlük oluşturur. Güncel derlemeler klinik potansiyeli vurgulamakla birlikte standartlaştırılmış, dıştan doğrulanmış ve iş akışına entegre modellerin sınırlı olduğunu belirtmektedir (Alabi vd., 2024).

Tablo 2: Tedavi yanıtı öngörüsünde seçilmiş radyomik kullanım alanları ve kanıt olgunluğu

Klinik bağlam	Görüntü	Başlıca görev	Kanıt olgunluğu	Öncelikli sınırlılık
KHDAK ve immünoterapi	BT, PET/BT	Yanıt ve sağkalım öngörüsü	Geliştirme / dış doğrulama sınırlı	Tedavi çizgisi, pseudoprogresyon ve merkez heterojenliği
Rektum kanseri ve neoadjuvan tedavi	T2 MRG, DWI/ADC	Patolojik tam yanıt	Çok sayıda retrospektif çalışma	Sekans, segmentasyon ve pCR tanımı farklılıkları
Meme kanseri ve neoadjuvan kemoterapi	DCE-MRG, DWI, PET	Patolojik yanıt	Sistematik derleme düzeyinde umut verici	Moleküler alt tip ve tedavi rejimi heterojenliği
Hepatoselüler karsinom ve TACE	Çok fazlı BT/MRG	Tedavi yanıtı ve risk sınıflaması	Erken dış doğrulama	Karaciğer fonksiyonu, teknik ve ölçüt farklılığı
Baş-boyun kanseri	BT, MRG, PET/BT	Yanıt, kontrol ve toksisite	Geliştirme aşaması	Artefakt, küçük hacim, HPV ve sonlanım çeşitliliği

Not. Kanıt olgunluğu; Chen vd. (2021), Liang vd. (2022), Feng vd. (2023), Alabi vd. (2024) ve Liao vd. (2025) temelinde nitel olarak özetlenmiştir; rutin klinik kullanım onayı anlamına gelmez.

MODEL GELİŞTİRME, DOĞRULAMA VE KLİNİK YARAR

Model geliştirme, algoritma seçiminden önce amaçlanan kullanımın tanımlanmasıyla başlamalıdır. Hangi hasta grubunda, hangi görüntüleme zamanında, hangi klinik kararı desteklemek üzere ve hangi alternatif bilgiye

ek olarak kullanılacağı belirtilmelidir. Hedef lezyon düzeyinde mi, hasta düzeyinde mi tahmin yapılacağı açık olmalıdır. Aynı hastaya ait birden fazla lezyon veya zaman noktası eğitim ve test kümelerine dağılırsa bağımsızlık varsayımı bozulur ve performans yapay olarak yükselir.

Radyomik veri kümelerinde değişken sayısı çoğu zaman hasta sayısından çok daha fazladır. Özellik seçimi, normalizasyon, imputasyon ve uyumlaştırma tüm veri üzerinde yapılırsa test kümesinden bilgi sızar. Doğru yaklaşım, bu adımların eğitim katları içinde öğrenildiği bir işlem hattıdır. Hiperparametre seçimi için iç döngü, performans tahmini için dış döngü kullanan iç içe çapraz doğrulama küçük veri kümelerinde daha gerçekçi bir iç doğrulama sağlar. Yine de çapraz doğrulama bağımsız dış doğrulamanın yerine geçmez (Vabalas vd., 2019).

Örneklem büyüklüğü yalnızca toplam hasta sayısı ile değil, sonuç olaylarının sayısı, tahmin değişkenlerinin etkin serbestlik derecesi ve beklenen model performansı ile değerlendirilmelidir. Çok sayıda aday özellik ve az sayıda yanıt olayı, özellik seçiminin kararsızlaşmasına ve farklı örneklemelerde bütünüyle farklı imzalar oluşmasına yol açabilir. Veri artırma görüntü çeşitliliğini destekleyebilir, ancak bağımsız hasta bilgisinin yerini tutmaz. Nadir sonlanımlarda sınıf ağırlığı veya yeniden örnekleme kullanılacaksa bu işlemler yalnızca eğitim verisinde uygulanmalı ve test kümesindeki gerçek prevalans korunmalıdır.

Referans standardının kalitesi, modelin ulaşabileceği performansın üst sınırını belirler. Patolojik tam yanıt için örnekleme ve tümör regresyon derecelendirmesi; RECIST için ölçüm yapan gözlemci, uzlaşma yöntemi ve değerlendirme zamanı; sağkalım için takip süresi ve sansürleme yeterliliği açıkça tanımlanmalıdır. Tek bir uzmanın etiketi mutlak gerçek olarak kabul edilmemeli, gözlemci belirsizliği ölçülmelidir. Algoritma ile uzman arasındaki uyumsuzluk her zaman algoritma hatası değildir; klinik açıdan önemli olgularda körlenmiş uzlaşma veya bağımsız kurul değerlendirmesi daha güçlü bir referans sağlayabilir.

Rastgele ayırma, aynı merkez ve zaman döneminden gelen benzer hastaları iki kümeye dağıttığı için dağılım kaymasına karşı iyimser olabilir. Zamansal doğrulama daha sonraki hastalarda, coğrafi doğrulama ise farklı kurumlarda performansı sınar. Dış doğrulama kohortu görüntüleme protokolü, hasta spektrumu ve sonuç prevalansı bakımından geliştirme kohortundan yeterince farklı olmalıdır. Bir modelin yalnızca dış veri

üzerinde yeniden ayarlanmasından sonraki başarısı ile herhangi bir ayarlama yapılmadan önceki taşınabilirliği ayrı verilmelidir.

Ayırt edicilik ölçütleri tek başına yeterli değildir. İkili sonuçlarda eğri altı alan, duyarlılık ve özgüllüğe ek olarak pozitif ve negatif öngörü değerleri klinik prevalansla birlikte sunulmalıdır. Olasılık üreten modellerde kalibrasyon eğrisi, kalibrasyon eğimi ve kesişimi; sağkalım modellerinde zaman bağımlı performans ve uygun kalibrasyon değerlendirilmelidir. Dengesiz veri kümelerinde doğruluk yanıltıcı olabilir. Eşik seçimi test verisine göre yapılmamalı, klinik sonuçların yanlış pozitif ve yanlış negatif maliyetleriyle ilişkilendirilmelidir.

Karar eğrisi analizi, bir modelin belirli eşik olasılıklarında herkese tedavi veya hiç kimseye tedavi stratejilerine göre net yararını inceleyebilir. Ancak kötü kalibre edilmiş bir modelde veya klinik olarak gerçekçi olmayan eşik aralığında karar eğrisi yanıltıcı olabilir. En güçlü kanıt, model destekli yaklaşımın mevcut bakıma kıyasla kararları, iş yükünü, gecikmeyi, toksisiteyi veya hasta sonuçlarını iyileştirip iyileştirmedigini prospektif etki çalışmasında göstermektir.

Raporlama ve yanlılık değerlendirme çerçeveleri, araştırmanın eksik yönlerini görünür kılar. CLEAR radyomik iş akışına, CLAIM 2024 tıbbi görüntülemeye yapay zekâ çalışmalarına, TRIPOD+AI klinik tahmin modellerinin raporlanmasına odaklanır; PROBAST+AI ise tahmin modellerinde yanlılık riski ve uygulanabilirliği değerlendirmek için güncellenmiş bir araç sunar (Kocak vd., 2023; Collins vd., 2024; Tejani vd., 2024; Moons vd., 2025). Sistematik kalite incelemeleri, radyomik çalışmalarında raporlama iyileşse de prospektif doğrulama ve klinik çeviri kanıtının hâlâ seyrek olduğunu göstermektedir (Barry vd., 2025).

KLİNİK UYGULAMAYA GEÇİŞ VE SINIRLILIKLAR

Bir modelin klinik kullanıma geçişindeki temel sorun, geliştirme verisindeki performansın yeni merkez, cihaz ve hasta gruplarında korunmasıdır. Alan kayması; tarayıcı, protokol, rekonstrüksiyon, kontrast uygulaması, hasta demografisi, hastalık evresi ve tedavi pratiğindeki değişikliklerden kaynaklanabilir. Modelin başarısızlığı yalnızca ortalama performans düşüşü olarak değil, belirli tümör tipleri veya hasta alt gruplarında sistematik hata olarak ortaya çıkabilir. Bu nedenle alt grup performansı yeterli örnekleme, önceden tanımlanmış biçimde incelenmelidir (Kelly vd., 2019).

Harmonizasyon ve yeniden kalibrasyon, dağılım farklarını azaltabilir; ancak sürekli model izleme gereksinimini ortadan kaldırmaz. Tarayıcı yazılım güncellemesi, yeni tedavi rejimi veya raporlama pratiğindeki değişim model girdilerini ve sonuç prevalansını değiştirebilir. Kullanım sonrasında veri kalitesi, eksik girdi oranı, performans, kalibrasyon ve güvenlik olayları izlenmeli; yeniden eğitim koşulları ve sürüm yönetimi önceden belirlenmelidir. Güncellenen modelin önceki sürümle karşılaştırılması ve sessiz doğrulama dönemi güvenli geçiş için yararlıdır.

Teknik entegrasyon klinik değer için zorunludur. PACS, radyoloji bilgi sistemi ve elektronik sağlık kaydıyla bütünleşmeyen bir araç ek veri aktarımı ve çift kayıt gerektirerek iş yükünü artırabilir. Çıktı, yalnızca bir risk puanı değil; ölçülen lezyonlar, önceki incelemeyle eşleşme, belirsizlik, veri kalitesi uyarıları ve modelin geçerli olduğu hasta grubuna ilişkin bilgi içermelidir. Kullanıcı arayüzü, radyoloğun düzeltme yapmasına ve bu düzeltmenin denetlenebilir biçimde kaydedilmesine izin vermelidir.

Canlı kullanımdan önce yürütülen sessiz doğrulamada model mevcut görüntüler üzerinde arka planda çalıştırılabilir; çıktısı tedavi kararına yansıtılmadan yerel performans, başarısızlık türleri ve iş akışı gereksinimleri ölçülür. Bu aşama, geliştirme çalışmasında görünmeyen DICOM etiket sorunlarını, seri seçimi hatalarını ve zaman aşımını ortaya çıkarabilir. Yerel eşik ayarı gerekiyorsa ayar kohortu ile nihai değerlendirme kohortu ayrılmalıdır. Sessiz doğrulamanın başarılı olması klinik etkiyi kanıtlamaz, fakat kontrolsüz devreye almaya göre önemli bir güvenlik basamağıdır.

İnsan denetimi, otomatik sistemin yalnızca son adımına eklenen formalite olmamalıdır. Hangi hataların kullanıcı tarafından fark edilebileceği, otomasyon yanlışlığı ve aşırı güven insan faktörleri çalışmalarıyla değerlendirilmelidir. Hekim ile algoritmanın tek başına performansı kadar birlikte performansı, okuma süresi ve anlaşmazlıkların çözümü de ölçülmelidir. Modelin hata yapabileceği örneklerin eğitim materyallerinde gösterilmesi, güvenli kullanım sınırlarının anlaşılmasını kolaylaştırır.

Federe öğrenme, ham görüntüleri merkez dışına çıkarmadan ortak model eğitimi olanağı sağlayabilir ve veri merkezileştirme gereksinimini azaltabilir. Buna karşın kurumlar arası altyapı, yönetim, etiket uyumu, veri dağılımı farklılığı, güvenli güncelleme ve olası bilgi sızıntısı sorunları devam eder. Dolayısıyla federe öğrenme veri paylaşımını ‘zorlaştıran’ bir yöntem değil, mahremiyet koruyucu iş birliği için bir seçenek; aynı zamanda yeni teknik ve yönetsel yükler getiren bir mimaridir (Rieke vd., 2020).

Etik, hukuki sorumluluk, mahremiyet ve düzenleyici değerlendirme kitabın ilk bölümünde ayrıntılı ele alınmıştır. Onkolojik görüntülemeye özgü olarak burada vurgulanması gereken nokta, model çıktısının tedavi kararını ve dolayısıyla toksisite ile hastalık kontrolünü etkileyebilmesidir. Amaçlanan kullanım, kullanıcı sorumluluğu, hata bildirimi, denetim kaydı ve performans değişikliği planı açık olmalıdır. Klinik kullanım öncesi yerel doğrulama yapılması, düzenleyici onayın her kurum ve hasta grubunda aynı performansı güvence altına aldığı varsayımından kaçınılması gerekir.

Maliyet-etkililik de klinik yararın parçasıdır. Otomatik ölçümün birkaç dakikalık zaman kazancı, yazılım lisansı, entegrasyon, kalite kontrol ve eğitim maliyetleriyle birlikte değerlendirilmelidir. Buna karşılık yüksek hacimli araştırma kohortlarında tutarlı ölçüm veya klinik araştırmalarda merkezi değerlendirme yükünün azaltılması daha erken ve ölçülebilir bir değer önerisi oluşturabilir. Rutin tedavi seçimi ise daha yüksek kanıt eşiği gerektirir.

SONUÇ VE ÖNERİLER

Radyomik ve yapay zekâ, onkolojik görüntülerde görünür boyut ölçümünün ötesindeki nicel örüntüleri kullanarak tedavi yanıtını öngörme ve değerlendirme potansiyeline sahiptir. En olgun uygulamalar, hedef lezyonların izlenmesi, ölçüm desteği ve büyük araştırma veri kümelerinde sonuç çıkarımı gibi sınırları belirli görevlerdir. Tedavi seçimini doğrudan değiştiren öngörü modellerinde ise kanıt çoğunlukla retrospektif geliştirme ve erken dış doğrulama aşamasındadır.

Klinik açıdan güvenilir bir çalışma; önceden tanımlanmış kullanım amacı, standart görüntüleme ve özellik çıkarımı, hasta düzeyinde veri ayrımı, eğitim içinde yürütülen ön işleme, uygun iç ve bağımsız dış doğrulama, kalibrasyon, klinik yarar analizi ve saydam raporlama içermelidir. Delta-radyomik ve çok modlu modeller, yalnızca daha karmaşık oldukları için üstün kabul edilmemeli; basit klinik ve görüntüleme modellerine karşı ek değerleri gösterilmelidir.

Gelecek çalışmaların odağı tek bir veri kümesinde en yüksek performansı elde etmekten, farklı merkezlerde tekrarlanabilirlik ve gerçek iş akışında hasta yararı göstermeye kaymalıdır. Yapay zekâ, radyoloğun klinik muhakemesinin yerine geçen bağımsız bir otorite olarak değil; ölçümü standartlaştıran, belirsizliği görünür kılan ve multidisipliner kararı destekleyen denetlenebilir bir sistem olarak konumlandırıldığında onkolojik görüntülemede kalıcı değer üretebilir.

KAYNAKÇA

- Abbas, E., Fanni, S. C., Bandini, C., Francischello, R., Febi, M., Aghakhanyan, G., Ambrosini, I., Faggioni, L., Cioni, D., Lencioni, R. A. ve Neri, E.. (2023). Delta-radiomics in cancer immunotherapy response prediction: A systematic review. *European Journal of Radiology Open*, 11, 100511. <https://doi.org/10.1016/j.ejro.2023.100511>
- Alabi, R. O., Elmusrati, M., Leivo, I., Almangush, A. ve Mäkitie, A. A.. (2024). Artificial Intelligence-Driven Radiomics in Head and Neck Cancer: Current Status and Future Prospects. *International Journal of Medical Informatics*, 188, 105464. <https://doi.org/10.1016/j.ijmedinf.2024.105464>
- Arbour, K. C., Luu, A. T., Luo, J., Rizvi, H., Plodkowski, A. J., Sakhi, M., Huang, K. B., Digumarthy, S. R., Ginsberg, M. S., Girshman, J., Kris, M. G., Riely, G. J., Yala, A., Gainor, J. F., Barzilay, R. ve Hellmann, M. D.. (2021). Deep Learning to Estimate RECIST in Patients with NSCLC Treated with PD-1 Blockade. *Cancer Discovery*, 11(1), 59-67. <https://doi.org/10.1158/2159-8290.CD-20-0419>
- Barry, N., Kendrick, J., Molin, K., Li, S., Rowshanfarzad, P., Hassan, G. M., Dowling, J., Parizel, P. M., Hofman, M. S. ve Ebert, M. A.. (2025). Evaluating the impact of the Radiomics Quality Score: a systematic review and meta-analysis. *European Radiology*, 35(3), 1701-1713. <https://doi.org/10.1007/s00330-024-11341-y>
- Bera, K., Braman, N., Gupta, A., Velcheti, V. ve Madabhushi, A.. (2022). Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nature Reviews Clinical Oncology*, 19(2), 132-146. <https://doi.org/10.1038/s41571-021-00560-7>
- Chen, Q., Zhang, L., Mo, X., You, J., Chen, L., Fang, J., Wang, F., Jin, Z., Zhang, B. ve Zhang, S.. (2021). Current status and quality of radiomic studies for predicting immunotherapy response and outcome in patients with non-small cell lung cancer: a systematic review and meta-analysis. *European Journal of Nuclear Medicine and Molecular Imaging*, 49(1), 345-360. <https://doi.org/10.1007/s00259-021-05509-7>
- Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A. L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., ... Logullo, P.. (2024). TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, e078378. <https://doi.org/10.1136/bmj-2023-078378>
- Dahm, I. C., Kolb, M., Altmann, S., Nikolaou, K., Gatidis, S., Othman, A. E., Hering, A., Moltz, J. H. ve Peisen, F.. (2024). Reliability of Automated RECIST 1.1 and Volumetric RECIST Target Lesion Response Evaluation in Follow-Up CT—A Multi-Center, Multi-Observer Reading Study. *Cancers*, 16(23), 4009. <https://doi.org/10.3390/cancers16234009>
- Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D. ve Verweij, J.. (2009). New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer*, 45(2), 228-247. <https://doi.org/10.1016/j.ejca.2008.10.026>
- Feng, L., Chen, Q., Huang, L. ve Long, L.. (2023). Radiomics features of computed tomography and magnetic resonance imaging for predicting response to transarterial chemoembolization in hepatocellular carcinoma: a meta-analysis. *Frontiers in Oncology*, 13, 1194200. <https://doi.org/10.3389/fonc.2023.1194200>
- Gillies, R. J., Kinahan, P. E. ve Hricak, H.. (2016). Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, 278(2), 563-577. <https://doi.org/10.1148/radiol.2015151169>
- Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G. ve King, D.. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17(1), 195. <https://doi.org/10.1186/s12916-019-1426-2>
- Kocak, B., Baessler, B., Bakas, S., Cuocolo, R., Fedorov, A., Maier-Hein, L., Mercaldo, N., Müller, H., Orlhac, F., Pinto dos Santos, D., Stanzione, A., Ugga, L. ve Zwanenburg, A.. (2023). CheckList for EvaluAtion of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII. *Insights into Imaging*, 14(1), 75. <https://doi.org/10.1186/s13244-023-01415-8>

- Lambin, P., Leijenaar, R. T. H., Deist, T. M., Peerlings, J., de Jong, E. E. C., van Timmeren, J., Sanduleanu, S., Larue, R. T. H. M., Even, A. J. G., Jochems, A., van Wijk, Y., Woodruff, H., van Soest, J., Lustberg, T., Roelofs, E., van Elmpt, W., Dekker, A., Mottaghy, F. M., Wildberger, J. E. ve Walsh, S.. (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nature Reviews Clinical Oncology*, 14(12), 749-762. <https://doi.org/10.1038/nrclinonc.2017.141>
- Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G. P. M., Granton, P., Zegers, C. M. L., Gillies, R., Boellard, R., Dekker, A. ve Aerts, H. J. W. L.. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*, 48(4), 441-446. <https://doi.org/10.1016/j.ejca.2011.11.036>
- Liang, X., Yu, X. ve Gao, T.. (2022). Machine learning with magnetic resonance imaging for prediction of response to neoadjuvant chemotherapy in breast cancer: A systematic review and meta-analysis. *European Journal of Radiology*, 150, 110247. <https://doi.org/10.1016/j.ejrad.2022.110247>
- Liao, Z., Luo, D., Tang, X., Huang, F. ve Zhang, X.. (2025). MRI-based radiomics for predicting pathological complete response after neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a systematic review and meta-analysis. *Frontiers in Oncology*, 15, 1550838. <https://doi.org/10.3389/fonc.2025.1550838>
- Liu, X., Wang, R., Zhu, Z., Wang, K., Gao, Y., Li, J., Zhang, Y., Wang, X., Zhang, X. ve Wang, X.. (2022). Automatic segmentation of hepatic metastases on DWI images based on a deep learning method: assessment of tumor treatment response according to the RECIST 1.1 criteria. *BMC Cancer*, 22(1), 1285. <https://doi.org/10.1186/s12885-022-10366-0>
- Liu, Z., Wang, S., Dong, D., Wei, J., Fang, C., Zhou, X., Sun, K., Li, L., Li, B., Wang, M. ve Tian, J.. (2019). The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics*, 9(5), 1303-1322. <https://doi.org/10.7150/thno.30309>
- Moons, K. G. M., Damen, J. A. A., Kaul, T., Hooft, L., Andaur Navarro, C., Dhiman, P., Beam, A. L., Van Calster, B., Celi, L. A., Denaxas, S., Denniston, A. K., Ghassemi, M., Heinze, G., Kengne, A. P., Maier-Hein, L., Liu, X., Logullo, P., McCradden, M. D., Liu, N., ... van Smeden, M.. (2025). PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*, 388, e082505. <https://doi.org/10.1136/bmj-2024-082505>
- Nardone, V., Reginelli, A., Rubini, D., Gagliardi, F., Del Tufo, S., Belfiore, M. P., Boldrini, L., Desideri, I. ve Cappabianca, S.. (2024). Delta radiomics: an updated systematic review. *La radiologia medica*, 129(8), 1197-1214. <https://doi.org/10.1007/s11547-024-01853-4>
- Nie, X., Chen, H., Gao, P., Yu, H., Sun, M. ve Yan, P.. (2024). Feasibility of an artificial intelligence system for tumor response evaluation. *BMC Medical Imaging*, 24(1), 280. <https://doi.org/10.1186/s12880-024-01460-9>
- Orlhac, F., Eertink, J. J., Cottreau, A. S., Zijlstra, J. M., Thieblemont, C., Meignan, M., Boellaard, R. ve Buvat, I.. (2022). A Guide to ComBat Harmonization of Imaging Biomarkers in Multicenter Studies. *Journal of Nuclear Medicine*, 63(2), 172-179. <https://doi.org/10.2967/jnumed.121.262464>
- Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R. M., Trask, A., Xu, D., Baust, M. ve Cardoso, M. J.. (2020). The future of digital health with federated learning. *npj Digital Medicine*, 3(1), 119. <https://doi.org/10.1038/s41746-020-00323-1>
- Scapicchio, C., Gabelloni, M., Barucci, A., Cioni, D., Saba, L. ve Neri, E.. (2021). A deep look into radiomics. *La radiologia medica*, 126(10), 1296-1311. <https://doi.org/10.1007/s11547-021-01389-x>
- Seymour, L., Bogaerts, J., Perrone, A., Ford, R., Schwartz, L. H., Mandrekas, S., Lin, N. U., Litière, S., Dancey, J., Chen, A., Hodi, F. S., Therasse, P., Hoekstra, O. S., Shankar, L. K., Wolchok, J. D., Ballinger, M., Caramella, C. ve de Vries, E. G. E.. (2017). iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *The Lancet Oncology*, 18(3), e143-e152. [https://doi.org/10.1016/S1470-2045\(17\)30074-8](https://doi.org/10.1016/S1470-2045(17)30074-8)

- Tareco Bucho, T. M., Petrychenko, L., Abdelatty, M. A., Bogveradze, N., Bodalal, Z., Beets-Tan, R. G. H. ve Trebeschi, S.. (2024). Reproducing RECIST lesion selection via machine learning: Insights into intra and inter-radiologist variation. *European Journal of Radiology Open*, 12, 100562. <https://doi.org/10.1016/j.ejro.2024.100562>
- Tejani, A. S., Klontzas, M. E., Gatti, A. A., Mongan, J. T., Moy, L., Park, S. H., Kahn, C. E. ve CLAIM 2024 Update Panel. (2024). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. *Radiology: Artificial Intelligence*, 6(4), e240300. <https://doi.org/10.1148/ryai.240300>
- Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A. T., Christian, M. C. ve Gwyther, S. G.. (2000). New Guidelines to Evaluate the Response to Treatment in Solid Tumors. *JNCI: Journal of the National Cancer Institute*, 92(3), 205-216. <https://doi.org/10.1093/jnci/92.3.205>
- Traverso, A., Wee, L., Dekker, A. ve Gillies, R.. (2018). Repeatability and Reproducibility of Radiomic Features: A Systematic Review. *International Journal of Radiation Oncology*Biophysics*, 102(4), 1143-1158. <https://doi.org/10.1016/j.ijrobp.2018.05.053>
- Vabalas, A., Gowen, E., Poliakoff, E. ve Casson, A. J.. (2019). Machine learning algorithm validation with a limited sample size. *PLOS ONE*, 14(11), e0224365. <https://doi.org/10.1371/journal.pone.0224365>
- van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J. C., Pieper, S. ve Aerts, H. J. W. L.. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, 77(21), e104-e107. <https://doi.org/10.1158/0008-5472.CAN-17-0339>
- Wahl, R. L., Jacene, H., Kasamon, Y. ve Lodge, M. A.. (2009). From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors. *Journal of Nuclear Medicine*, 50(Suppl 1), 122S-150S. <https://doi.org/10.2967/jnumed.108.057307>
- Woo, M., Devane, A. M., Lowe, S. C., Lowther, E. L. ve Gimbel, R. W.. (2021). Deep learning for semi-automated unidirectional measurement of lung tumor size in CT. *Cancer Imaging*, 21(1), 43. <https://doi.org/10.1186/s40644-021-00413-7>
- Zwanenburg, A., Vallières, M., Abdalah, M. A., Aerts, H. J. W. L., Andrearczyk, V., Apte, A., Ashrafinia, S., Bakas, S., Beukinga, R. J., Boellaard, R., Bogowicz, M., Boldrini, L., Buvat, I., Cook, G. J. R., Davatzikos, C., Depeursinge, A., Desseroit, M. C., Dinapoli, N., Dinh, C. V., ... Löck, S.. (2020). The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*, 295(2), 328-338. <https://doi.org/10.1148/radiol.2020191145>

GİRİŞİMSSEL RADYOLOJİDE YAPAY ZEKA: İŞLEM PLANLAMA VE NAVİGASYON TEKNOLOJİLERİ

İbrahim Feyyaz Naldemir*

GİRİŞ

Girişimsel radyoloji, görüntüleme yöntemlerini tanı ve tedavi amacıyla birleştiren; teknik başarı kadar hasta seçimi, anatomik planlama, gerçek zamanlı karar verme ve işlem sonrası izlemin de belirleyici olduğu bir disiplindir. Floroskopi, ultrasonografi (US), bilgisayarlı tomografi (BT), manyetik rezonans görüntüleme (MRG) ve konik ışınlı BT (cone-beam CT, CBCT) ile üretilen görüntüler; klinik, laboratuvar ve cihaz verileriyle birlikte değerlendirildiğinde çok katmanlı bir karar ortamı oluşur. Yapay zeka (YZ), bu veri akışından örüntü çıkarma, anatomik yapıları bölütleme, risk tahmini yapma, hedef veya cihaz konumunu izleme ve görüntü kalitesini iyileştirme gibi görevlerde girişimsel radyoloğa destek olma potansiyeline sahiptir (Glielmo vd., 2024; Lesaunier vd., 2025).

Bu potansiyel, girişimsel radyolojinin tamamen otonom bir alana dönüştüğü anlamına gelmez. Güncel uygulamaların önemli bölümü retrospektif model geliştirme, prototip, fantom, simülasyon veya sınırlı klinik doğrulama aşamasındadır. Ticari bir navigasyon yazılımının, otomatik görüntü kaydının ya da uzaktan kumandalı robotun kullanılması da tek başına YZ kullanıldığı anlamına gelmez. Teknolojinin niteliği; girdinin ne olduğu, çıktının nasıl üretildiği, sistemin hangi kararı etkilediği ve hekimin müdahale düzeyi açıklanarak tanımlanmalıdır (O'Brien vd., 2024; Huth vd., 2025).

Girişimsel radyolojide YZ için en uygun değerlendirme çerçevesi işlem öncesi, işlem sırası ve işlem sonrası evreleri birlikte ele almaktır. İşlem öncesinde hasta seçimi, risk katmanlaması, otomatik bölütleme ve yol planlama; işlem sırasında görüntü füzyonu, hedef ve cihaz takibi, perfüzyon analizi ile doz optimizasyonu; işlem sonrasında ise teknik başarının nicel değerlendirilmesi, ablasyon sınır analizi, komplikasyon ve yeniden girişim

* Dr. Öğr. Üyesi, Düzce Üniversitesi, Tıp Fakültesi, Radyoloji Anabilim Dalı, Düzce, dr.ifnaldemir@gmail.com,
ORCID: <https://orcid.org/0000-0002-0804-7378>

riskinin öngörülmesi öne çıkar. Her evrede kullanılacak sonlanım ölçütleri farklıdır ve bir modelin yüksek doğruluk göstermesi, işlem süresini veya hasta sonucunu iyileştirdiğini tek başına kanıtlamaz.

Bu bölümde YZ, otomasyon ve robotik kavramları birbirinden ayrılarak girişimsel radyolojideki başlıca kullanım alanları incelenmekte; mevcut kanıtın olgunluğu, klinik doğrulama gereksinimleri ve güvenli uygulamanın temel koşulları tartışılmaktadır. Amaç, teknolojik olasılıkları abartılı bir otonomi anlatısıyla değil, ölçülebilir klinik görevler ve insan gözetimi temelinde değerlendirmektir.

GİRİŞİMSSEL RADYOLOJİDE YAPAY ZEKANIN KAVRAMSAL ÇERÇEVESİ

Yapay Zeka, Otomasyon ve Navigasyonun Ayrımı

Kural tabanlı otomasyon, önceden tanımlanmış eşik veya geometrik kuralları uygular; makine öğrenmesi ise örnek verilerden bir ilişki öğrenerek yeni veride tahmin üretir. Derin öğrenme, özellikle görüntü bölütleme, nesne saptama, görüntü kaydı ve gürültü azaltmada kullanılan çok katmanlı modelleri kapsar. Bir sistemin otomatik çalışması, onun mutlaka YZ içerdiğini göstermez. Benzer biçimde, YZ tabanlı bir karar destek sistemi motor bir görevi yerine getirmeyebilir. Bu ayrım, hem kanıtın yorumlanması hem de cihazın güvenlik ve düzenleyici değerlendirmesi açısından önemlidir (Iezzi vd., 2019; Lesaunier vd., 2025).

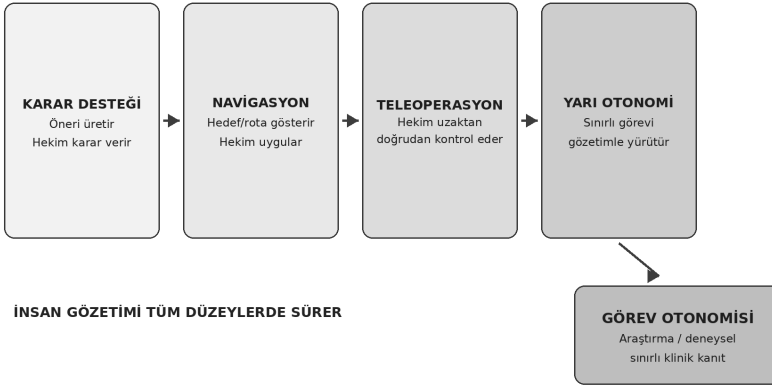
Navigasyon, hedef ile cihaz arasındaki uzamsal ilişkiyi belirlemeye yardımcı olan geniş bir kavramdır. Elektromanyetik izleme, optik izleme, füzyon görüntüleme, CBCT tabanlı yol haritalama ve robotik kateter kontrolü navigasyon başlığı altında yer alabilir. Bu sistemlerin bazıları yalnızca geometrik eşleştirme yaparken bazıları organ deformasyonunu, hedef hareketini veya cihaz konumunu öğrenilmiş modellerle tahmin eder. Bu nedenle bilimsel metinlerde kullanılan sistemin YZ bileşeni açıkça tarif edilmeli; üretici tarafından kullanılan “akıllı” ifadesi bilimsel sınıflandırma yerine geçmemelidir.

Robotik ve Otonomi Düzeyleri

Robotik sistemlerde teleoperasyon, hekimin kateteri, kılavuz teli veya iğneyi uzaktaki bir konsoldan doğrudan yönlendirmesidir. Hareket ölçekleme ve titreme filtreleme gibi özellikler teknik hassasiyeti artırabilir; ancak karar ve hareket komutu operatöre aittir. Yarı otonom sistemler, önceden belirlenmiş bir yol üzerinde ilerleme veya iğneyi tanımlı bir hedefe

hizalama gibi sınırlı görevleri gözetim altında yerine getirir. Görev otonomisinde sistem belirli bir işlemsel alt görevi planlayıp yürütebilir; tam otonomi ise hasta seçimi, strateji, uygulama ve komplikasyon yönetimini insan müdahalesi olmadan kapsar. Son iki düzey, girişimsel radyolojide ağırlıklı olarak araştırma alanındadır (Ning vd., 2023; Robertshaw vd., 2024).

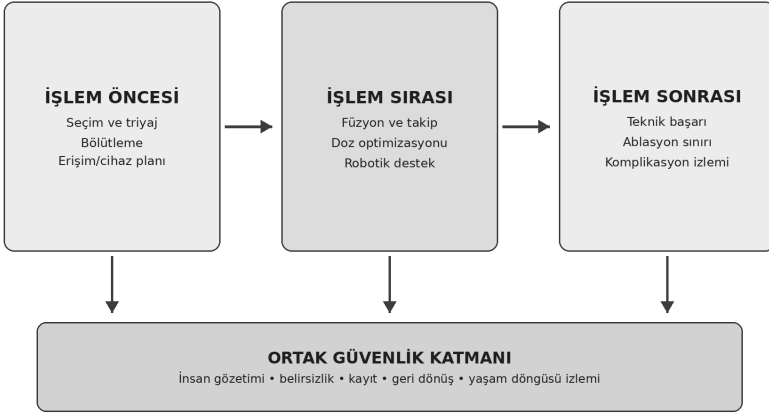
Otonomi düzeyi yükseldikçe yalnızca teknik doğruluk değil, hata saptama, güvenli duruş, manuel kontrole geçiş, gecikme süresi, siber güvenlik ve operatörün durumsal farkındalığı da doğrulanmalıdır. Sistem, anatomik değişkenlik veya beklenmeyen direnç karşısında güvenilir biçimde sınırlarını tanıyamıyorsa otonom hareket üstün bir teknik özellik değil, yeni bir risk kaynağı olur.



Şekil 2: Karar desteğinden deneysel görev otonomisine uzanan teknoloji ve otonomi spektrumu (yazar tarafından hazırlanmıştır).

İnsan–Yapay Zeka İş Birliği

Klinik hedef, radyoloğun yerini alan bir sistem geliştirmekten çok, tekrarlayan ve ölçülebilir görevleri destekleyerek bilişsel yükü azaltmaktır. YZ; aday hedefleri işaretleyebilir, alternatif erişim yollarını sıralayabilir veya beklenen doz ve işlem süresi hakkında tahmin sunabilir. Girişimsel radyolog ise önerinin klinik bağlama uygunluğunu, tedavi alternatiflerini, hastanın tercihlerini ve beklenmeyen durumları değerlendirir. İnsan–YZ iş birliğinin başarısı yalnızca model performansına değil, bilginin ekranda nasıl sunulduğuna, belirsizliğin görünür olup olmadığına ve yanlış güvenin önlenmesine bağlıdır (Huth vd., 2025).



Şekil 1: Girişimsel radyolojide işlem öncesi, işlem sırası ve işlem sonrası yapay zeka uygulamaları ile ortak güvenlik katmanı (yazar tarafından hazırlanmıştır).

İŞLEM ÖNCESİ UYGULAMALAR

Hasta Seçimi, Triyaj ve Risk Katmanlaması

İşlem öncesi karar desteğinin ilk hedefi, doğru hastayı doğru zamanda doğru tedaviye yönlendirmektir. Model girdileri demografik özellikler, komorbiditeler, laboratuvar sonuçları, önceki tedaviler, rapor metinleri ve görüntü tabanlı özellikleri içerebilir. Çıktı; teknik başarısızlık, komplikasyon, uzun yatış, yeniden girişim veya belirli bir tedaviden yarar olasılığı olabilir. Bununla birlikte tahmin hedefi tedavi kararından önce tanımlanmalı; modelin klinik karar verilirken gerçekten mevcut olmayan bilgileri kullanmasına izin verilmemelidir.

Akut iskemik inmede büyük damar oklüzyonunun otomatik saptanması ve ilgili ekibin eş zamanlı bilgilendirilmesi, zaman duyarlı iş akışına örnek oluşturur. Bu sistemler görüntülerin önceliklendirilmesini, merkezler arası iletişimi ve sevk kararını hızlandırabilir. Ancak bildirim süresindeki iyileşme ile fonksiyonel sonucun iyileşmesi aynı sonlanım değildir; klinik yarar, trombektomiye erişim, kapı-kasık zamanı, yeniden kanalizasyon ve fonksiyonel sonuçlar ayrı ayrı değerlendirilmelidir (Al-Janabi vd., 2024).

Abdominal aort anevrizmasında YZ tabanlı yaklaşımlar çap ölçümüne ek olarak anevrizma geometrisi, intraluminal trombüs, duvar özellikleri ve biyomekanik parametreleri birleştirerek büyüme veya rüptür riskini araştırmaktadır. Endovasküler aort onarımı planlamasında otomatik merkez hattı çıkarımı, boyun ve iliak erişim ölçümleri cihaz seçimini destekleyebilir.

Bununla birlikte tek bir merkezde geliştirilen yüksek performanslı modelin farklı cihaz ve popülasyonlarda aynı sonucu vereceği varsayılmamalıdır (Raffort vd., 2020).

Otomatik Bölütleme ve Üç Boyutlu Vasküler Haritalama

Karaciğer, böbrek, akciğer ve damar sisteminin otomatik bölütlenmesi; hedef, risk altındaki yapı ve güvenli erişim koridorunun üç boyutlu olarak gösterilmesini sağlar. Derin öğrenme modelleri portal ve hepatik venleri, hepatik arter dallarını, tümörleri ve komşu organları ayrı sınıflar halinde tanımlayabilir. Bu çıktı transarteriyel kemoembolizasyon (TAKE), transarteriyel radyoembolizasyon (TARE), portal ven embolizasyonu ve perkütan ablasyon gibi işlemlerde planlamayı destekleyebilir (Chierici vd., 2024).

Vasküler varyasyonların sıklığı ve küçük damarların kesit kalınlığı, kontrast zamanlaması ve hareketten etkilenmesi, otomatik haritalamanın temel güçlükleridir. Hepatik arter anatomisinin çok sayıda varyant içermesi nedeniyle model çıktısı kaynak görüntülerle doğrulanmalıdır (Ugurel vd., 2010). Özellikle “besleyici arter” etiketi, yalnızca anatomik devamlılığa değil tümör perfüzyonuna ilişkin bir çıkarım içerir. Yanlış pozitif bir dal gereksiz kateterizasyona, yanlış negatif sonuç ise yetersiz tedaviye yol açabilir.

Tümör ve damar bölütleme performansı Dice benzerlik katsayısı veya yüzey mesafesiyle raporlanabilir; ancak girişimsel kullanımda aşağı akım sonuç daha önemlidir. Örneğin birkaç milimetrelilik sınır hatası, geniş bir organda yüksek Dice değerine rağmen iğne yolunu veya ablasyon sınırını klinik olarak değiştirebilir. Bu nedenle geometrik metrikler, hedefe erişim hatası, düzeltme gereksinimi ve plan değişikliği gibi görev temelli metriklerle birlikte değerlendirilmelidir.

Erişim Yolu, Cihaz ve Tedavi Planlaması

Perkütan işlemlerde planlama sistemi hedefe ulaşan aday yolları üretebilir; plevra, bağırsak, safra yolu, büyük damar veya sinir gibi kaçınılması gereken yapılarla mesafeyi hesaplayabilir. Termal ablasyonda tümörü kapsayacak prob sayısı, yerleşim açısı ve beklenen ablasyon hacmi tahmin edilebilir. Endovasküler işlemlerde damar çapı, kıvrımlılık, kalsifikasyon ve darlık yükü kateter, kılavuz tel, stent-greft veya embolik ajan seçimine yardımcı olabilir. Bu çıktılar öneri niteliğinde olmalı ve girişimsel radyoloğun uygulanabilirlik değerlendirmesinden geçmelidir.

Karaciğer tümörlerinde süperselektif TAKE için kullanılan bazı rehberlik yazılımları CBCT verisinden hedef ve olası besleyici arterleri göstererek kateterizasyonu destekler. Miyayama ve arkadaşlarının çalışması, küçük hepatoselüler karsinomlarda rehberlik yazılımıyla yürütülen süperselektif konvansiyonel TAKE'nin klinik uygulanabilirliğini göstermektedir; ancak kullanılan yazılımın her bileşeni YZ olarak sınıflandırılmamalıdır (Miyayama vd., 2021). Bu örnek, klinik yarar gösteren navigasyon ile öğrenen algoritma kavramlarının ayrı tutulması gerektiğini vurgular.

Planlama modelleri yalnızca anatomik “en kısa yolu” değil, klinik olarak uygulanabilir yolu optimize etmelidir. Solunum hareketi, hastanın pozisyonu, prob açısı, steril alan, gantri açıklığı ve olası hidrodiseksiyon gibi işlemsel kısıtlar eğitim verisine veya optimizasyon fonksiyonuna yansıtılmadığında öneri teorik olarak doğru fakat pratikte kullanılamaz olabilir.

Tablo 1: Girişimsel radyolojide yapay zeka uygulamalarının işlem evresine göre sınıflandırılması

Evre	Başlıca görev	Uygun çıktı/sonlanım	Kanıtın genel olgunluğu
İşlem öncesi	Triyaj, risk tahmini, bölütleme, damar ve erişim planlama	Zaman, plan değişikliği, hedef/rota doğruluğu	Retrospektif çalışmalar yaygın; harici ve prospektif doğrulama sınırlı
İşlem sırası	Füzyon, hedef-cihaz takibi, CBCT/DSA analizi	Hedefleme hatası, teknik başarı, işlem süresi	Fantom ve fizibilite çalışmaları baskın; göreve göre klinik kanıt değişken
Doz yönetimi	Düşük doz görüntü iyileştirme veya sentezi	Hava kerması, DAP, görüntü kalitesi, komplikasyon	Üretken YZ tabanlı DSA için çok merkezli randomize kanıt mevcut
Robotik	Teleoperasyon, paylaşımlı kontrol, sınırlı görev otonomisi	Hedefe ulaşma, kuvvet, süre, manuel dönüş, güvenli duruş	Teleoperasyon klinik uygulanabilir; otonom navigasyon çoğunlukla deneysel
İşlem sonrası	Ablasyon sınırı, tedavi dağılımı, komplikasyon/yeniden girişim	En küçük sınır, ek tedavi, lokal kontrol, alarm yükü	Teknik doğrulama geliyor; hasta sonucuna etkisi için prospektif çalışma gerekli

Not: DAP, doz-alan çarpımı; DSA, dijital subtraksiyon anjiyografisi; CBCT, konik ışınlı BT. Olgunluk düzeyi, uygulamalar arasında farklılık gösterebilir.

İş Akışı ve Kaynak Planlaması

YZ tabanlı modeller, planlanan işlemin süresini, anestezi gereksinimini, özel sarf malzemesi veya ek ekip ihtiyacını öngörerek oda çizelgelemesini destekleyebilir. Bu kullanım alanında hedef, yalnızca daha fazla işlem planlamak değil; gecikme, iptal, fazla mesai ve acil olgular için gerekli

kapasiteyi dengeli biçimde yönetmektir. Modelin başarısı tahmin hatasının yanı sıra hasta bekleme süresi, oda kullanım oranı, son dakika iptali ve personel yükündeki değişimle değerlendirilmelidir (Huth vd., 2025; Lesaunier vd., 2025).

Operasyonel model, geçmişte belirli hasta gruplarına daha uzun süre ayrılmış olmasını “verimsizlik” olarak öğrenebilir ve mevcut eşitsizlikleri sürdürebilir. Karmaşık hastaların daha kısa zaman dilimlerine zorlanması güvenliği azaltabileceğinden, çizelgeleme çıktısı klinik karmaşıklık ve kaynak kısıtlarıyla birlikte insan denetiminden geçmelidir.

İŞLEM SIRASI UYGULAMALAR

Görüntü Kaydı ve Füzyon

Görüntü füzyonu, işlem öncesi yüksek çözünürlüklü görüntü ile işlem sırasındaki US, floroskopi veya CBCT'nin ortak koordinat sisteminde birleştirilmesidir. Rijit kayıt kemik yapılar veya sınırlı hareket gösteren bölgelerde yeterli olabilirken; karaciğer, akciğer ve böbrekte solunum, hasta pozisyonu, iğne basısı, hidrodiseksiyon veya ablasyona bağlı deformasyon nedeniyle esnek kayıt gerekebilir. YZ tabanlı kayıt modelleri bu dönüşümü hızlandırmayı amaçlar; ancak hız, doğruluğun ve güvenilirlik denetiminin yerine geçmez (Glielmo vd., 2024).

Füzyon hatası hedefin yakınında ölçülmeli ve kullanıcıya görünür olmalıdır. Tüm organ düzeyinde iyi görünen çakışma, küçük bir lezyon çevresinde klinik açıdan kabul edilemez hata içerebilir. Sistem, kayıt güveni düşük olduğunda uyarı vermeli; yeniden kayıt, ek CBCT veya alternatif görüntüleme gereksinimini operatöre bırakmalıdır. Özellikle yalnızca sanal hedefe bakılarak ilerlenen işlemlerde kaynak görüntünün aralıklı doğrulanması zorunludur.

Hedef, İğne ve Kateter Takibi

US kılavuzluğunda iğne görünürlüğü açısı, doku yapısı, gölgelenme ve operatör hareketinden etkilenir. Evrişimsel sinir ağları iğne gövdesi ve ucunu gerçek zamanlı görüntüde saptayarak beklenen yörüngeyi gösterebilir. Fantom ve deneysel çalışmalarda gerçek zamanlı iğne saptama uygulanabilir bulunmuş; ancak klinik kullanım için farklı prob, iğne, derinlik ve artefakt koşullarında doğrulama gereksinimi devam etmiştir (Mwikirize vd., 2018; Arapi vd., 2023).

Kateter ve kılavuz tel takibi, floroskopik görüntüde cihazın segmentasyonu ile damar yol haritasının birleştirilmesine dayanabilir. Bir sistemin cihazı doğru saptaması tek başına güvenli navigasyon anlamına gelmez; damar duvarına uygulanan kuvvet, kateter ucunun görüntü düzlemi dışındaki konumu ve ilerlemeye karşı direnç gibi bilgiler de önemlidir. Görüntü temelli izleme, kuvvet algılama ve robotik kontrol bir araya getirildiğinde yeni güvenlik olanakları kadar yeni hata zincirleri de oluşur.

İşlem Sırası Görüntü Analizi

CBCT ve dijital subtraksiyon anjiyografisi, embolizasyon sırasında hedef perfüzyonu ve tedavi dağılımını değerlendirmek için kullanılabilir. Otomatik damar bölütleme, tümör boyanmasının nicelleştirilmesi ve lipiodol birikiminin üç boyutlu analizi, sonlanımın yalnızca görsel değerlendirmeye bırakılmasını azaltabilir. Derin öğrenmeyle CBCT üzerindeki lipiodol birikiminin otomatik bölütlenmesi araştırılmıştır; bu tür sistemlerin gerçek klinik değeri, segmentasyon başarısının yanı sıra ek embolizasyon kararını ve lokal tümör kontrolünü değiştirip değiştirmediğiyle değerlendirilmelidir (Malpani vd., 2022).

Perfüzyon ve embolizasyon sonlanımı için evrensel bir eşik yoktur. Aşırı agresif sonlanım non-hedef iskemi riskini, yetersiz sonlanım ise rezidüel canlı tümör riskini artırabilir. YZ çıktısının yalnızca renkli bir harita olarak sunulması yerine, belirsizlik, önceki enjeksiyonla değişim ve kateter konumuyla ilişki gibi bilgilerle desteklenmesi daha güvenli bir karar ortamı sağlar.

Radyasyon ve Kontrast Optimizasyonu

Doz optimizasyonunda YZ; düşük doz görüntülerde gürültü azaltma, yüksek kaliteli görüntü sentezi, otomatik kolimasyon önerisi veya görüntü gereksiniminin tahmini için kullanılabilir. Bu uygulamalar, bakış takibi ya da otomatik puls frekansı değişikliği ile özdeş değildir; sistemin gerçek mekanizması açıkça belirtilmelidir. Zhao ve arkadaşlarının çok merkezli randomize çalışmasında üretken YZ tabanlı düşük doz DSA sistemi, standart protokole kıyasla hava kerması ve doz-alan çarpımını yaklaşık üçte iki oranında azaltmış; işlem süresi ve komplikasyonlar bakımından önceden tanımlı karşılaştırmalar yapılmıştır (Zhao vd., 2026). Bu çalışma klinik düzeyi yüksek bir kanıt sunmakla birlikte, sonuçların cihaz, işlem türü ve uygulama protokolü dışındaki ortamlara doğrudan genellenmemesi gerekir.

Kontrast optimizasyonu için model, önceki enjeksiyonlar, böbrek fonksiyonu, hemodinami ve hedef damar bilgisine dayalı öneri üretebilir. Ancak görüntü kalitesini koruyan en düşük kontrast hacmi; kateter konumu, akım ve cihaz parametreleriyle yakından ilişkilidir. Bu alanda hasta güvenliğini gösteren prospektif çalışma sayısı, teknik fizibilite çalışmalarından daha azdır.

Robotik Navigasyon ve Deneysel Otonomi

Uzaktan kumandalı robotik sistemler, operatörün radyasyon alanı dışında çalışmasına, hareketi ölçeklendirmesine ve kateter-kılavuz tel manevralarını konsoldan yönetmesine olanak tanır. Robotik yardımcı karotis anjiyoplasti ve stentleme için yayımlanan teknik rehberler, teleoperasyonun uygulanabilirliğini göstermektedir; bunlar kendi başına otonom YZ kanıtı değildir (Moskalik vd., 2026). Klinik değerlendirmede teknik başarıya ek olarak konsola geçiş süresi, manuel dönüş gereksinimi, cihaz uyumluluğu, radyasyon maruziyeti ve komplikasyonlar raporlanmalıdır.

Robertshaw ve arkadaşları mekanik trombektomi bağlamında ters pekiştirmeli öğrenmeyle kateter ve kılavuz tel navigasyonunu araştırmıştır. Çalışma, sınırlı bir görev için otonom navigasyonun kavram kanıtını sunar; hastalarda hekimden üstün veya hatasız bir uygulamayı kanıtlamaz (Robertshaw vd., 2024). Otonom sistemlerde başarı yalnızca hedefe ulaşma oranıyla değil; damar duvarı teması, uygulanan kuvvet, yanlış dala giriş, süre, kurtarma manevrası ve güvenli duruş davranışıyla ölçülmelidir.

Klinik uygulamaya en yakın yaklaşım, tam otonomiden çok paylaşımlı kontroldür. Sistem aday yolu veya sınırlı hareketi önerirken hekim onay verir, ilerlemeyi izler ve her an kontrolü devralabilir. Paylaşımlı kontrolün başarısı, hekimin pasif gözlemciye dönüşmemesine ve otomasyon arızasını zamanında tanıyabilmesine bağlıdır.

İŞLEM SONRASI UYGULAMALAR

Anlık Teknik Başarı ve Ablasyon Sınırı

Termal ablasyonda amaç, tümörü tümüyle kapsayan ve hastalık türüne uygun güvenlik sınırı içeren bir ablasyon hacmi oluşturmaktır. Gerekli sınır tümör biyolojisi, organ, yöntem ve komşu kritik yapılara göre değiştiğinden tüm olgular için değişmez bir “5 mm kuralı” olarak sunulmamalıdır. Karaciğer tümörlerinde yeterli ablasyon sınırı yerel kontrol açısından önemlidir ve standart uygulama belgelerinde tedavi planının temel bileşeni olarak ele alınır (Crocetti vd., 2020).

İşlem öncesi tümör hacmi ile işlem sonrası ablasyon hacminin üç boyutlu kaydı, en küçük ablasyon sınırını ve yetersiz kapsanan alanı gösterebilir. Derin öğrenme tabanlı esnek görüntü kaydı, termal deformasyon ve solunuma bağlı farkları telafi etmeyi amaçlar. Hepatoselüler karsinomda mikrodalga ablasyon sonrası sınır değerlendirmesi ve kolorektal karaciğer metastazlarında otomatik bölütleme-kayıt yaklaşımları umut verici sonuçlar vermiştir; ancak algoritmik sınır ölçümünün yeniden ablasyon kararı ve uzun dönem lokal kontrol üzerindeki etkisi prospektif olarak gösterilmelidir (An vd., 2020; Lin vd., 2023).

Ablasyon planlama yazılımları beklenen ablasyon hacmini üretici verisi veya öğrenilmiş modelle tahmin edebilir. Gerçek hacim; doku perfüzyonu, prob gücü, uygulama süresi, ısı kaybı ve komşu damarlar nedeniyle farklılaşabilir. Bu nedenle tahmin, işlem sonrası görüntülemenin yerine geçmemeli; sapma model performansının izlenmesinde kullanılmalıdır (Westby vd., 2025).

Embolizasyon Sonrası Değerlendirme

TAKE veya TARE sonrasında teknik başarı; hedef damarların seçimi, tedavi dağılımı, non-hedef embolizasyon ve rezidüel tümör perfüzyonu birlikte değerlendirilerek belirlenir. YZ, seri CBCT görüntülerini eşleştirebilir, lipiodol veya kontrast tutulumunu nicelleştirebilir ve tedavi edilmeyen tümör bölümlerini işaretleyebilir. Bu çıktılar işlem masasında ek tedavi kararı için kullanılacaksa yanlış pozitif ve yanlış negatif sonuçların klinik maliyeti ayrıca analiz edilmelidir.

İşlem sonrası erken görüntüleme bulgusuyla haftalar sonraki tedavi yanıtı aynı kavram değildir. Teknik sonlanım, hedef bölgenin işlem sırasında planlandığı şekilde tedavi edildiğini; klinik etkinlik ise takipte tümör kontrolü veya hasta yararı sağlandığını gösterir. YZ araştırmalarında bu sonlanımlar birbirinin yerine kullanılmamalıdır.

Komplikasyon, Yeniden Girişim ve İzlem

İşlem sonrası vital bulgular, laboratuvar sonuçları, hemşire gözlemleri ve görüntüleme verileri birlikte kullanılarak kanama, akut böbrek hasarı, enfeksiyon, pnömotoraks veya uzamış yatış riski tahmin edilebilir. Mevcut çalışmaların önemli bölümü retrospektiftir ve işlem öncesi risk tahminine odaklanır (Glielmo vd., 2024). Modelin saatler öncesinden komplikasyonu kesin olarak saptadığı veya mortaliteyi azalttığı, bu sonucu gösteren prospektif etki çalışması bulunmadan ileri sürülmemelidir.

Erken uyarı sisteminde duyarlılık kadar alarm yükü de önemlidir. Çok sayıda yanlış alarm, kullanıcıların uyarıları görmezden gelmesine ve gerçek riskin kaçırılmasına yol açabilir. Klinik değerlendirmede uyarı başına pozitif öngörü değeri, hasta başına alarm sayısı, uyarıdan müdahaleye kadar geçen süre ve bakım sürecindeki gerçek değişiklik birlikte raporlanmalıdır.

MODEL DOĞRULAMA VE KLİNİK UYGULAMAYA GEÇİŞ

Kullanım Amacı ve Sonlanım Ölçütleri

Klinik değerlendirme, “YZ doğruluğu” gibi genel bir hedef yerine kullanım amacının açık tanımıyla başlar. Sistem hangi hasta grubunda, hangi aşamada, hangi kullanıcı tarafından ve hangi kararı desteklemek için kullanılacaktır? Örneğin besleyici arter saptama modeli ile düşük doz DSA rekonstrüksiyon modeli aynı başarı ölçütleriyle değerlendirilemez. İlkinde lezyon veya damar düzeyinde duyarlılık, yanlış dal önerisi ve kateterizasyon süresi; ikincisinde görüntü kalitesi, hava kerması, doz-alan çarpımı ve klinik olarak önemli bulguların korunması ölçülmelidir.

Girişimsel radyolojide görev temelli sonlanımlar arasında teknik başarı, hedefleme hatası, işlem ve floroskopi süresi, hava kerması, doz-alan çarpımı, kontrast hacmi, cihaz değişimi, manuel müdahale, komplikasyon, yeniden girişim ve yatış süresi yer alır. Modelin ayırım performansı için duyarlılık, özgüllük veya eğri altında kalan alanın yanı sıra kalibrasyon, güven aralığı ve karar eşliğine bağlı klinik yarar raporlanmalıdır.

Doğrulama Tasarımı ve Genellenebilirlik

İç doğrulama, geliştirme verisindeki aşırı uyumu azaltmak için gereklidir; klinik genellenebilirliği kanıtlamaz. Harici doğrulama farklı merkez, cihaz, protokol ve hasta dağılımında yapılmalı; mümkünse zamansal ve coğrafi olarak bağımsız veri kullanılmalıdır. Girişimsel verilerde aynı hastaya ait çok sayıda görüntünün eğitim ve test kümelerine dağılması veri sızıntısına yol açabileceğinden bölme işlemi hasta düzeyinde yapılmalıdır (Eche vd., 2021).

Çalışma akışı, kayıp veriler, dışlama ölçütleri, referans standart, körleme, başarısız otomasyon ve insan düzeltmeleri saydam biçimde raporlanmalıdır. CLAIM 2024, tıbbi görüntüleme YZ çalışmalarının raporlanması için veri, model, değerlendirme ve şeffaflık başlıklarını güncellemiştir (Tejani vd., 2024). Canlı klinik ortamda erken aşama değerlendirme için DECIDE-AI; kullanılabilirlik, insan-sistem etkileşimi, performansın bağlam içinde

değerlendirilmesi ve güvenlik olaylarının raporlanmasına yönelik bir çerçeve sunar (Vasey vd., 2022).

Alt grup performansı yaş, cinsiyet, vücut yapısı, hastalık şiddeti, cihaz üreticisi ve görüntü kalitesine göre incelenmelidir. Nadir anatomik varyantlar, metal artefaktı, hareket ve düşük kontrast gibi zor olgular yalnızca dışlanmamalı; sistemin yanıt vermekten kaçınma oranı ve güvenli yönlendirme davranışı ayrıca raporlanmalıdır.

Belirsizlik tahmini, yalnızca olasılık değerini ekranda göstermekten ibaret değildir. Model güveninin gerçek hata olasılığıyla uyumlu olup olmadığı kalibrasyonla değerlendirilmeli; güvenin düşük olduğu durumlarda sistemin çıktı vermemesi veya insan incelemesine yönlendirmesi önceden tanımlanmalıdır. Bu seçici davranışın başarısı, yalnızca yanıt verilen olgulardaki doğrulukla değil, tüm olgular içindeki kaçınma oranı ve kaçınılan olguların klinik özellikleriyle birlikte raporlanmalıdır.

İş Akışı Entegrasyonu ve İnsan Faktörleri

Klinik olarak doğru bir model, iş akışına uygun değilse yarar üretmeyebilir. DICOM ve PACS entegrasyonu, sonuç üretme süresi, veri aktarım kesintileri, kullanıcı arayüzü ve denetim kaydı uygulamanın temel parçalarıdır. Gerçek zamanlı navigasyonda birkaç saniyelik gecikme kabul edilemezken işlem öncesi planlamada daha uzun süre tolere edilebilir. Ekranda sunulan bilgi, görüntünün önemli bölümünü örtmemeli ve kullanıcıya çıktıyı kaynak veride doğrulama olanağı sağlamalıdır.

Otomasyon yanlılığı, kullanıcının sistem önerisini yeterince sorgulamadan kabul etmesidir. Bunun karşıtı olan otomasyon isteksizliği ise tek bir hatadan sonra yararlı sistemi tümüyle reddetmektir. Eğitim, yalnızca yazılımın nasıl açılacağını değil; hangi koşullarda başarısız olabileceğini, belirsizliğin nasıl yorumlanacağını ve manuel iş akışına nasıl dönüleceğini kapsamalıdır. FUTURE-AI çerçevesi, güvenilir klinik YZ için adalet, evrensellik, izlenebilirlik, kullanılabilirlik, sağlamlık ve açıklanabilirlik ilkelerini birlikte ele almaktadır (Lekadir vd., 2025).

Klinik etki değerlendirmesine ekonomik ve operasyonel sonuçlar da eklenmelidir. Lisans, donanım, entegrasyon, bakım ve eğitim maliyetleri; kazanılan işlem süresi, azalan görüntü tekrarı, kontrast ve sarf tüketimi ya da önlenen komplikasyonlarla birlikte incelenmelidir. Zaman kazandırdığı gösterilen bir sistemin toplam maliyeti artırması veya uzman düzeltmesine

bağımlı kalması mümkündür. Bu nedenle maliyet-etkililik, yalnızca üretici tarafından bildirilen iş akışı süresine dayanarak varsayılmamalıdır.

Tablo 2: Klinik uygulamaya geçiş için asgari doğrulama ve güvenlik bileşenleri

Bileşen	Asgari değerlendirme	Girişimsel radyolojiye özgü örnek	Yetersiz değerlendirme riski
Kullanım amacı	Hasta, kullanıcı, zaman noktası, hedef karar ve çıktı	Besleyici arter önerisi veya düşük doz DSA	Aynı modelin farklı kararlarda kontrolsüz kullanılması
Harici doğrulama	Farklı merkez, cihaz, protokol ve hasta grubu	Farklı anjiyografi cihazı ve embolizasyon protokolü	Merkez dışı performans kaybı
Görev temelli performans	Teknik ve klinik sonuçları, güven aralıkları	Hedefleme hatası, floroskopi süresi, kontrast hacmi	Yüksek görüntü metriğine rağmen klinik yetersizlik
İnsan faktörleri	Düzeltilme süresi, alarm yükü, otomasyon yanlılığı	Yanlış damar önerisinin operatörce fark edilmesi	Yanlış güven veya uyarı yorgunluğu
Arıza ve geri dönüş	Güvenli duruş, manuel kontrol, kurtarma süresi	Robotik navigasyondan manuel kateterizasyona geçiş	Gecikme veya doğrudan hasta zararı
Yaşam döngüsü	Sürüm, veri kayması, performans ve olay izlemi	Yeni cihaz yazılımı sonrası yeniden doğrulama	Sessiz performans bozulması

Not: Bu bileşenler CLAIM, DECIDE-AI ve FUTURE-AI ilkeleriyle uyumlu biçimde, sistemin risk düzeyine göre ayrıntılandırılmalıdır.

Dağıtım Sonrası İzleme

Model performansı klinik kullanıma alındıktan sonra sabit kalmayabilir. Yeni cihaz yazılımı, protokol değişikliği, farklı embolik materyal, kullanıcı alışkanlığı veya hasta dağılımındaki değişim veri kaymasına yol açabilir. Bu nedenle performans göstergeleri sürüm bazında izlenmeli; eşik altına düşüldüğünde sistemin sınırlandırılması veya yeniden doğrulanması için önceden tanımlı plan bulunmalıdır.

Güncelleme süreci, önceki sürümle karşılaştırma, geriye dönük uyumluluk, kullanıcı bilgilendirmesi ve değişiklik kaydını içermelidir. Sürekli öğrenen bir sistemde hangi verinin ne zaman modele katıldığı ve yeni sürümün hangi klinik koşullarda onaylandığı izlenebilir olmalıdır.

GÜVENLİK, ETİK VE HUKUKİ ÇERÇEVE

İnsan Gözetimi ve Güvenli Geri Dönüş

Girişimsel radyolojide YZ çıktısı, doğrudan invaziv bir eyleme dönüşebileceği için tanınan karar desteğine göre farklı bir risk profiline

sahiptir. Hatalı bir lezyon işareti gereksiz biyopsiye, yanlış damar seçimi non-hedef embolizasyona, hatalı kayıt ise organ yaralanmasına yol açabilir. İnsan gözetimi, yalnızca son ekrandaki “onay” düğmesine basmak değildir; kaynak veriyi inceleyebilen, öneriyi değiştirebilen ve sistemi durdurabilen yetkin bir operatörü gerektirir.

Her sistem için güvenli geri dönüş yöntemi tanımlanmalıdır. Ağ bağlantısı, görüntü aktarımı, kayıt veya robotik kontrol kesildiğinde işlem manuel olarak sürdürülebilmeli ya da güvenli biçimde sonlandırılmalıdır. Acil durdurma işlevi, kontrol devri ve kurtarma süresi simülasyonla test edilmelidir. Cihaz arızalarının ve yakın olayların düzenli kaydı, performans izleminin ayrılmaz parçasıdır.

Veri Güvenliği, Açıklanabilirlik ve Sorumluluk

Görüntüler, cihaz telemetrisi ve klinik verilerin ortak kullanımı veri güvenliği riskini artırır. Kimliksizleştirme, erişim yetkisi, şifreleme, kayıt tutma ve siber saldırıya karşı dayanıklılık sistem tasarımında ele alınmalıdır. Robotik veya gerçek zamanlı sistemde veri bütünlüğünün bozulması yalnızca mahremiyet sorunu değil, doğrudan hasta güvenliği sorunudur.

Açıklanabilirlik gereksinimi kullanım amacına göre değişir. Görüntü iyileştirme sisteminde radyoloğun asıl gereksinimi klinik bulgunun korunması ve yapay ayrıntı üretilmediğinin gösterilmesiyken; tedavi öneren modelde hangi değişkenlerin kararı etkilediği daha önemlidir. Görsel dikkat haritası tek başına nedensel açıklama değildir ve yanlış güven oluşturabilir (Brady ve Neri, 2020).

Hukuki sorumluluk ülkeye, ürünün düzenleyici statüsüne, kullanım amacına, üretici talimatına ve hekimin sistemi nasıl kullandığına göre değişebilir. Bu nedenle tüm sorumluluğun evrensel olarak yalnızca operatörde olduğu şeklinde kategorik bir ifade doğru değildir. Kurum; yetkilendirme, eğitim, sürüm takibi, olay bildirimi ve yerel yönetişim süreçlerini tanımlamalı, hekim ise sistemi onaylanmış kullanım amacı ve klinik standartlar içinde kullanmalıdır.

Girişimsel Radyoloğun Değişen Rolü

YZ ve robotik, girişimsel radyoloğun anatomik bilgi, el becerisi ve komplikasyon yönetimi sorumluluğunu ortadan kaldırmaz. Buna veri okuryazarlığı, model sınırlarını tanıma, kalite göstergelerini izleme ve insan-sistem etkileşimini yönetme görevlerini ekler. Gelecekte bazı tekrarlayan alt görevler otomatikleşebilir; fakat tedavi hedefinin belirlenmesi, alternatiflerin

hasta ile paylaşılması, beklenmeyen duruma uyum ve etik sorumluluk klinisyenin temel rolü olarak kalacaktır.

SONUÇ VE ÖNERİLER

Güvenli uygulama için sistemin kullanım amacı ve YZ bileşeni açıkça tanımlanmalı; klinik, deneysel ve simülasyon kanıtı birbirinden ayrılmalıdır. Harici ve prospektif doğrulama, görev temelli sonlanımlar, başarısız otomasyon oranı, insan düzeltmesi, alarm yükü, alt grup performansı ve dağıtım sonrası izleme temel değerlendirme ölçütleridir. Robotik uygulamalarda manuel kontrol, güvenli duruş ve kurtarma senaryoları teknik başarının ayrılmaz parçasıdır.

Yakın dönemde en gerçekçi model, girişimsel radyoloğun kontrolünü koruduğu, YZ'nin aday hedef ve yollar sunduğu, tekrarlayan ölçümleri otomatikleştirdiği ve belirsiz durumda geri çekildiği paylaşımlı karar ve kontrol yaklaşımıdır. Teknik uygulanabilirlik, klinik yarar ve otonomi aynı kavramlar değildir; klinik değer, teknolojinin ne kadar “otonom” görüldüğüyle değil, doğru hastada güvenliği, tutarlılığı ve kaynak kullanımını ölçülebilir biçimde iyileştirmesiyle belirlenmelidir.

KAYNAKÇA

- Al-Janabi, O. M., El Refa'ei, A., Elgazzar, T., Mahmood, Y. M., Bakir, D., Gajjar, A., ... Brinjikji, W. (2024). Current Stroke Solutions Using Artificial Intelligence: A Review of the Literature. *Brain Sciences*, 14(12), 1182. <https://doi.org/10.3390/brainsci14121182>
- An, C., Jiang, Y., Huang, Z., Gu, Y., Zhang, T., Ma, L. ve Huang, J. (2020). Assessment of Ablative Margin After Microwave Ablation for Hepatocellular Carcinoma Using Deep Learning-Based Deformable Image Registration. *Frontiers in Oncology*, 10, 573316. <https://doi.org/10.3389/fonc.2020.573316>
- Arapi, V., Hardt-Stremayr, A., Weiss, S. ve Steinbrener, J. (2023). Bridging the simulation-to-real gap for AI-based needle and target detection in robot-assisted ultrasound-guided interventions. *European Radiology Experimental*, 7(1), 30. <https://doi.org/10.1186/s41747-023-00344-x>
- Brady, A. P. ve Neri, E. (2020). Artificial Intelligence in Radiology—Ethical Considerations. *Diagnostics*, 10(4), 231. <https://doi.org/10.3390/diagnostics10040231>
- Chierici, A., Lareyre, F., Salucki, B., Iannelli, A., Delingette, H. ve Raffort, J. (2024). Vascular liver segmentation: a narrative review on methods and new insights brought by artificial intelligence. *Journal of International Medical Research*, 52(9), 03000605241263170. <https://doi.org/10.1177/03000605241263170>
- Crocetti, L., de Baére, T., Pereira, P. L. ve Tarantino, F. P. (2020). CIRSE Standards of Practice on Thermal Ablation of Liver Tumours. *CardioVascular and Interventional Radiology*, 43(7), 951-962. <https://doi.org/10.1007/s00270-020-02471-z>
- Eche, T., Schwartz, L. H., Mokrane, F. Z. ve Dercle, L. (2021). Toward Generalizability in the Deployment of Artificial Intelligence in Radiology: Role of Computation Stress Testing to Overcome Underspecification. *Radiology: Artificial Intelligence*, 3(6), e210097. <https://doi.org/10.1148/ryai.2021210097>
- Glielmo, P., Fusco, S., Gitto, S., Zantonelli, G., Albano, D., Messina, C., ... Mauri, G. (2024). Artificial intelligence in interventional radiology: state of the art. *European Radiology Experimental*, 8(1), 62. <https://doi.org/10.1186/s41747-024-00452-2>
- Huth, H., Saccenti, L., Anibal, J., Ukeh, I., Kassim, M., Gichoya, J., ... Wood, B. J. (2025). Artificial Intelligence and Intelligent Interventional Radiology: Fusion of Human and Machine. *Journal of Vascular and Interventional Radiology*, 36(12), 1929-1936.e1. <https://doi.org/10.1016/j.jvir.2025.09.017>
- Iezzi, R., Goldberg, S. N., Merlino, B., Posa, A., Valentini, V. ve Manfredi, R. (2019). Artificial Intelligence in Interventional Radiology: A Literature Review and Future Perspectives. *Journal of Oncology*, 2019, 1-5. <https://doi.org/10.1155/2019/6153041>
- Lekadir, K., Frangi, A. F., Porras, A. R., Glocker, B., Cintas, C., Langlotz, C. P., ... Starmans, M. P. A. (2025). FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ*, 388, e081554. <https://doi.org/10.1136/bmj-2024-081554>
- Lesaunier, A., Khlaoui, J., Dancette, C., Tselikas, L., Bonnet, B. ve Boeken, T. (2025). Artificial intelligence in interventional radiology: Current concepts and future trends. *Diagnostic and Interventional Imaging*, 106(1), 5-10. <https://doi.org/10.1016/j.diii.2024.08.004>
- Lin, Y. M., Paolucci, I., O'Connor, C. S., Anderson, B. M., Rigaud, B., Fellman, B. M., ... Odisio, B. C. (2023). Ablative Margins of Colorectal Liver Metastases Using Deformable CT Image Registration and Autosegmentation. *Radiology*, 307(2), e221373. <https://doi.org/10.1148/radiol.221373>
- Malpani, R., Petty, C. W., Yang, J., Bhatt, N., Zeevi, T., Chockalingam, V., ... Lin, M. (2022). Quantitative Automated Segmentation of Lipiodol Deposits on Cone-Beam CT Imaging Acquired during Transarterial Chemoembolization for Liver Tumors: A Deep Learning Approach. *Journal of Vascular and Interventional Radiology*, 33(3), 324-332.e2. <https://doi.org/10.1016/j.jvir.2021.12.017>
- Miyayama, S., Yamashiro, M., Ikeda, R., Matsumoto, J., Takeuchi, K., Sakuragawa, N., ... Terada, T. (2021). Efficacy of Superselective Conventional Transarterial Chemoembolization Using Guidance Software for Hepatocellular Carcinoma within Three Lesions Smaller Than 3 cm. *Cancers*, 13(24), 6370. <https://doi.org/10.3390/cancers13246370>

- Moskalik, A. D., Liang, B., Cord, B. J. ve Waldau, B. (2026). A technical guide to robotic-assisted carotid angioplasty and stenting with the Corindus CorPath GRX system. *Interventional Neuroradiology*, 32(4), 2066-2066. <https://doi.org/10.1177/15910199241226856>
- Mwirikize, C., Noshier, J. L. ve Hacıhaliloglu, I. (2018). Convolution neural networks for real-time needle detection and localization in 2D ultrasound. *International Journal of Computer Assisted Radiology and Surgery*, 13(5), 647-657. <https://doi.org/10.1007/s11548-018-1721-y>
- Ning, S., Chautems, C., Kim, Y., Rice, H., Hanning, U., Al Kasab, S., ... Abdalkader, M. (2023). Robotic Interventional Neuroradiology: Progress, Challenges, and Future Prospects. *Seminars in Neurology*, 43(03), 432-438. <https://doi.org/10.1055/s-0043-1771298>
- O'Brien, A. J., Vrazas, J. I. ve Clements, W. (2024). Current applications of algorithmic artificial intelligence in interventional radiology: A review of the literature. *Journal of Medical Imaging and Radiation Oncology*, 68(2), 194-207. <https://doi.org/10.1111/1754-9485.13609>
- Raffort, J., Adam, C., Carrier, M., Ballaith, A., Coscas, R., Jean-Baptiste, E., ... Lareyre, F. (2020). Artificial intelligence in abdominal aortic aneurysm. *Journal of Vascular Surgery*, 72(1), 321-333.e1. <https://doi.org/10.1016/j.jvs.2019.12.026>
- Robertshaw, H., Karstensen, L., Jackson, B., Granados, A. ve Booth, T. C. (2024). Autonomous navigation of catheters and guidewires in mechanical thrombectomy using inverse reinforcement learning. *International Journal of Computer Assisted Radiology and Surgery*, 19(8), 1569-1578. <https://doi.org/10.1007/s11548-024-03208-w>
- Tejani, A. S., Klontzas, M. E., Gatti, A. A., Mongan, J. T., Moy, L., Park, S. H., ... Zins, M. (2024). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): 2024 Update. *Radiology: Artificial Intelligence*, 6(4), e240300. <https://doi.org/10.1148/ryai.240300>
- Ugurel, M. S., Battal, B., Bozlar, U., Nural, M. S., Tasar, M., Ors, F., ... Karademir, I. (2010). Anatomical variations of hepatic arterial system, coeliac trunk and renal arteries: an analysis with multidetector CT angiography. *The British Journal of Radiology*, 83(992), 661-667. <https://doi.org/10.1259/bjr/21236482>
- Vasey, B., Nagendran, M., Campbell, B., Clifton, D. A., Collins, G. S., Denaxas, S., ... Perkins, Z. B. (2022). Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nature Medicine*, 28(5), 924-933. <https://doi.org/10.1038/s41591-022-01772-9>
- Westby, K., Westby, D., McKevitt, K. ve Moloney, B. M. (2025). Artificial Intelligence in Thermal Ablation: Current Applications and Future Directions in Microwave Technologies. *Biomimetics*, 10(12), 818. <https://doi.org/10.3390/biomimetics10120818>
- Zhao, H., Bai, Y., Chen, L., Ma, J., Lei, Y., Sun, T., ... Zheng, C. (2026). Generative AI-based low-dose digital subtraction angiography for intra-operative radiation dose reduction: a randomized controlled trial. *Nature Medicine*, 32(1), 288-296. <https://doi.org/10.1038/s41591-025-04042-6>