

Mamografi ve Dijital Meme Tomosentezinde Yapay Zekânın Rolü

Yeliz Başar¹

Özet

Meme kanseri, dünya genelinde kadınlarda kansere bağlı ölümlerin başlıca nedenlerinden biri olmaya devam etmektedir. Bu nedenle mamografik tarama programları erken tanı açısından büyük önem taşımaktadır. Bununla birlikte, konvansiyonel mamografi taramalarında bildirilen mortalite azalmasına rağmen, kanserlerin yaklaşık %20–40'ının gözden kaçabildiği bildirilmiştir. Son yıllarda dijital meme tomosentezinin (DMT) kullanıma girmesi, kanser saptama oranlarını artırmış ve geri çağırma oranlarını azaltmıştır. Ancak DMT'nin hacimsel görüntüleme yapısı, radyologların inceleme süresini %50–200 oranında artırarak tarama programlarında önemli bir iş yükü oluşturmuştur.

Birinci nesil bilgisayar destekli tanı (CADE) sistemleri klinik uygulamada beklenen tanısal faydayı sağlayamamış ve yüksek yanlış pozitif oranları nedeniyle güvenilirlik sorunları yaratmıştır. Bu durum, özellikle DMT incelemelerinin çok sayıda kesit içeren yapısıyla birleştiğinde, radyologların mevcut iş yükünü daha da artırmaktadır.

Son yıllarda geliştirilen derin öğrenme tabanlı yapay zekâ sistemleri, mamografide kanser tespitinde radyolog performansına ulaşabilmekte hatta bazı çalışmalarda bu performansı aşabilmektedir. Yapay zekânın eş zamanlı karar destek sistemi olarak kullanılması, radyologların tanısal doğruluğunu artırırken okuma süresini artırmadan tanı performansını iyileştirebilmektedir. DMT taramalarında yapılan çalışmalar, klinik olarak anlamlı duyarlılık kaybı olmadan %30–70 oranında iş yükü azalmasının mümkün olduğunu göstermektedir. MASAI randomize çalışması, yapay zekâ destekli tek okuma yaklaşımının standart çift okumaya kıyasla daha fazla kanser tespit ettiğini ve radyolog iş yükünü belirgin biçimde azalttığını ortaya koymuştur.

1 Department of Radiology ,Acıbadem Maslak Hospital, Istanbul, Turkey, ORCID: 0000-0001-5562-0635

Bununla birlikte, algoritmaların farklı popülasyonlarda genellenebilirliği, bazı lezyon tiplerinde sınırlı duyarlılık, algoritmik önyargı ve medicolegal sorumluluk gibi konular halen araştırılmaya devam etmektedir.

1. Giriş

Meme kanseri, 2022 yılında dünya genelinde teşhis edilen yaklaşık 2,3 milyon yeni vaka ve 670.000 ölüm ile kadınlarda en sık görülen kanserlerden biri olup önemli bir halk sağlığı sorunu oluşturmaktadır. Bu yük, kanserlerin daha erken ve daha tedavi edilebilir evrelerde tespit edilmesini amaçlayan toplum tabanlı tarama programlarının geliştirilmesini ve yaygınlaştırılmasını teşvik etmiştir. (Siegel vd., 2024; Jemal vd., 2011). Bu hastalığın halk sağlığı üzerindeki yükü, kanserleri daha erken ve daha tedavi edilebilir evrelerde tespit etmek gibi kapsayıcı bir hedefle on yıllar boyunca toplum tabanlı tarama stratejilerine yatırım yapılmasını teşvik etmiştir.

Mamografik tarama, toplum düzeyinde meme kanseri mortalitesini azalttığı kanıtlanmış tek görüntüleme modalitesi olmaya devam etmektedir. Dönüm noktası niteliğindeki İsveç İki İlçe Çalışması mamografik taramaya atfedilebilir %30'luk bir mortalite düşüşünü ortaya koymuş (Tabar vd., 1985) ve bu fayda daha sonra ardışık tarama turlarıyla kümülatif mortalite düşüşünü gösteren prospektif modelleme çalışmalarıyla doğrulanıp genişletilmiştir (Duffy vd., 2021).

Söz konusu kanıtlanmış etkinliğe rağmen, mamografik taramanın doğasında bazı kusurlar barındırmaktadır. Meme kanserlerinin %20 ila %40'ı, düzenli bir tarama incelemesi ile bir sonraki planlanmış tarama arasındaki dönemde teşhis edilen tümörler olan "interval (ara dönem) kanserler" olarak ortaya çıkmaktadır. Önceki görüntülerin retrospektif incelemesi, bu lezyonların önemli bir kısmının önceki taramada görünür olduğunu, ancak tespit edilemediğini ortaya koymaktadır (Bird vd., 1992; Kızıldağ Yirgin vd., 2022). Bu hataların kaynakları arasında üst üste binen normal parankimden kaynaklanan doku örtüşmesi (maskeleme), radyolog yorgunluğu ve arama tatmini (search satisfaction) ile ilişkili algısal hatalar ve yüksek hacimli tarama pratiğinden doğan bilişsel kısıtlılıklar yer almaktadır.

Dijital meme tomosentezi (DMT), geleneksel iki boyutlu tam alanlı dijital mamografiye (FFDM) (Full-Field Digital Mammography) özgü doku örtüşmesi sorununu çözmek için geliştirilmiştir. Sınırlı bir açılarda birden fazla düşük dozlu X-ışını projeksiyonu elde ederek ve meme hacmini bir dizi ince kesitsel dilim olarak yeniden yapılandırarak, DMT süperpozisyon artefaktlarını büyük ölçüde ortadan kaldırır. Büyük ölçekli klinik çalışmalar, DMT'nin FFDM'ye kıyasla kanser tespit oranlarını yaklaşık %30-90 oranında

artırdığını ve yanlış pozitif geri çağırma oranlarını anlamlı ölçüde azalttığını doğrulamaktadır (Ciatto vd., 2013; Haas vd., 2013; Friedewald vd., 2014). Bununla birlikte, bir DMT incelemesini okumak, geleneksel 2B mamografiye göre %50-200 daha fazla zaman gerektirir –tipik bir çekim görünüm başına 40-200 ince rekonstrükte dilim üretir– ki bu durum, yüksek hacimli tarama pratiğinin mevcut yükünü katlayan bilişsel olarak zorlayıcı bir görevdir (Tagliafico vd., 2017).

Bu nedenle modern meme görüntülemesinin karşı karşıya kaldığı temel zorluk iki boyutludur: Bir yandan hacim ve yorgunluk koşulları altında insan algısının doğasında var olan kısıtlılıklara rağmen gözden kaçırma oranları nasıl en aza indirilecektir; diğer yandan DMT'nin artan yorumlama talebi, buna paralel olarak genişlemeyen iş gücü ve kaynak kısıtlamaları içerisinde nasıl karşılanacaktır. Yapay zekâ (yapay zekâ) –ve özelde derin öğrenme paradigması– her iki soruna da hedefe yönelik çözümler sunmaktadır.

Radyolojide bilgisayar destekli görüntü analizinin kavramsal kökenleri dijital çağa dayanmaktadır. Meyers vd. (1964), 1960'ların ortalarında radyografik görüntülerin otomatik bilgisayar analizini tanımlamış ve Spiesberger (1979) bu on yılın sonuna gelindiğinde mamogram incelemesi için hesaplamalı bir yaklaşım önermiştir. Modern yapay zekâyâ geçişte asıl dönüm noktası 2012 yılında, Krizhevsky vd.'nin evrişimli sinir ağının (CNN) –AlexNet– ImageNet büyük ölçekli görsel tanıma yarışmasında önceki tüm yöntemlerin çok altında hata oranlarına ulaştığını göstererek, tıbbi görüntüleme de dahil olmak üzere vizyon tabanlı disiplinlerde derin öğrenmenin (derin öğrenme) benimsenmesini hızlandırmasıyla gerçekleşmiştir.

Bu bölüm, birinci nesil bilgisayar destekli tespitten (CAD) çağdaş derin öğrenme tabanlı yapay zekâyâ evrimi izlemekte; bir dizi klinik uygulama genelinde mamografi ve DMT'de yapay zekâ için mevcut kanıt temelini gözden geçirmekte; Türkiye toplum tabanlı tarama programlarından elde edilen kanıtlar dâhil olmak üzere uygulama çerçevelerini ve gerçek dünya dağıtım kanıtlarını incelemekte; ve geniş klinik dağıtımla birlikte ele alınması gereken etik, düzenleyici ve yönetim hususlarını ele almaktadır.

2. CAD'den Derin Öğrenmeye: Teknik ve Tarihsel Evrim

2.1 Birinci Nesil Bilgisayar Destekli Tespit

Bilgisayar destekli tespit (CADE) ve bilgisayar destekli tanı (CADx) sistemleri, otomatik görüntü analizi üzerine on yıllar süren akademik araştırmaların ardından 1990'larda ticari ürünler olarak ortaya çıkmıştır. Birinci nesil CAD sistemleri geleneksel makine öğrenimi (MÖ) metodolojilerine dayanıyordu:

kural tabanlı ön işleme algoritmalarıyla tanımlanan ilgi alanlarından (ROI) form metrikleri, doku istatistikleri ve kenar gradyan analizleri dahil olmak üzere tasarlanmış özellik tanımlayıcıları (feature descriptors) hesaplanıyor; ardından bir sınıflandırıcı (genellikle destek vektör makineleri, rastgele ormanlar veya manuel tasarlanmış girdilere sahip yapay sinir ağları) iyi huylu bulguları, kötü huylu bulguları ve normal dokuyu ayırt etmek üzere eğitiliyordu. Mikro kalsifikasyon kümeleri, kitleler ve yapısal distorsiyonlar gibi her bulgu türü için ayrı işleme boru hatları gerekiyordu. Bu durum, bu mimarilerin ham piksel verilerinden genelleştirilmiş temsiller öğrenme konusundaki yetersizliğini yansıtıyordu.

Amerika Birleşik Devletleri Gıda ve İlaç Dairesi (FDA) ilk mamografi CADe sistemini 1998'de onayladı ve Medicare ve Medicaid Hizmetleri Merkezleri (CMS) 2002'de yapay zekâ destekli mamografi yorumlaması için geri ödeme kapsamını genişletti. Düzenleyici onayın ve finansal teşvikin bu kombinasyonu, kısmen CAD işaretlerini belgelenmiş ikincil bir kontrol olarak konumlandıran medikolegal baskıların da motive etmesiyle Amerikan tarama uygulamaları genelinde hızlı bir benimsemeye yol açtı. CAD'in fiili tanınal katkısına yönelik daha büyük şüpheleri yansıtan Avrupa'daki benimsenme süreci ise daha temkinliydi.

Bu şüphencilğin haklı olduğu ortaya çıktı. Birinci nesil CAD sistemleri, kanser tespitinde orantılı bir iyileşme sağlamadan inceleme başına 0,5 ila 2,5 arasında sahte (spurious) işaretleme ürettiyordu. Meme Kanseri Gözetim Konsorsiyumu'ndan (Breast Cancer Surveillance Consortium) elde edilen büyük ölçekli kanıtlar, CAD'in genel mamografi performansını artırmadığını ve kanser tespitinde karşılık gelen kazanımlar olmaksızın geri çağırma oranlarında istatistiksel olarak anlamlı artışlarla ilişkili olduğunu göstermiştir (Lehman vd., 2015). Dönüm noktası niteliğindeki bir çalışma, CAD kullanımının artan geri çağırma oranları ve biyopsi önerilerinin pozitif prediktif değerinde bir azalma ile ilişkili olduğunu bulmuştur (Fenton vd., 2007). Toplu kanıtlar, nihayetinde 2018'de CAD için CMS geri ödemesinin geri çekilmesine yol açmıştır. Aşağıdaki Tablo 1, geleneksel CAD ile günümüzün derin öğrenme tabanlı yapay zekâsı arasındaki temel farklılıkları özetlemektedir.

Tablo 1. Birinci Nesil CAD ve Derin Öğrenme yapay zekâ — Temel Farklılıklar

| Özellik | Geleneksel CAD (1990'lar-2010'lar) | Derin Öğrenme yapay zekâ (2015-günümüz) |
|------------------------------|--|--|
| Özellik Mühendisliği | Alan uzmanları tarafından el ile şekillendirildi; şekil, doku ve kenar tanımlayıcılarının açıkça belirtilmesini gerektirdi. | Geri yayılım (backpropagation) yoluyla ham piksel verilerinden katman katman otomatik olarak öğrenilir; manuel özellik tanımlaması gerekmez. |
| Eğitim Yaklaşımı | Gözetimli MÖ: SVM'ler, rastgele ormanlar, elle tasarlanmış özellik vektörleri üzerinde artırma (boosting). | Geniş etiketlenmiş görüntü veri setlerinde uçtan uca (end-to-end) CNN eğitimi (örn. >45.000 inceleme; Kooi vd., 2017). |
| Yanlış Pozitif Oranı | Yüksek: İnceleme başına 0,5-2,5 sahte işaretleme; klinik kabulün önündeki büyük engel. | Büyük ölçüde azaltılmış; bağlamsal hiyerarşik özellik öğrenimi sahte işaretlemeleri baskılar. |
| AUC (EAA) Performansı | Genellikle 0,70-0,80; ortalama radyolog performansının altında. | Büyük ölçekli doğrulamalarda 0,84-0,93; ortalama radyolog performansında veya üzerinde (Rodríguez-Ruiz vd., 2019b; Park vd., 2024). |
| FDA Onay Yılı | 1998 (ilk CADe sistemi, ImageChecker, R2 Technology). | 2020 (DMT için Transpara); 2018'den günümüze çok sayıda derin öğrenme sistemi onaylandı. |
| Okuma Süresine Etkisi | Net bir azalma yok; yanlış pozitif işaretleri gözden geçirme yükü nedeniyle ılımlı artış. | Nötr ile mütevazı azalma arasında; yapay zekâ güdümlü dikkat yönlendirmesi DMT'de arama süresini kısaltır (Van Winkel vd., 2021). |
| Mevcut Klinik Durum | Aşamalı olarak kaldırılmaktadır; CMS geri ödemesi 2018'de geri çekildi; net zarar kanıtı (Fenton vd., 2007; Lehman vd., 2015). | Dünya çapındaki organize tarama programlarında aktif olarak kullanılmaktadır; düzenleyici onaylar genişlemektedir. |

(Şekil 1. Birinci nesil CAD (el yapımı özellik veri yolu → sınıflandırıcı) ile derin öğrenme CNN'sinin (evrişim ve havuzlama katmanları aracılığıyla ham piksellerden uçtan uca özellik öğrenimi ile sınıflandırma çıktısı) şematik karşılaştırması.)

2.2 Derin Öğrenme Devrimi

Krizhevsky vd.'nin 2012 ImageNet yarışması sonucu, vizyon tabanlı alanlarda modern yapay zekânın dönüm noktası olarak geniş çapta kabul görmektedir. AlexNet'in CNN'i %15,3'lük bir ilk-5 (top-5) hata oranına ulaşarak –en iyi ikinci rakibi için bu oran %26,2 idi– CNN'lerin el yapımı

özellik mühendisliği olmaksızın, doğrudan verilerden görsel içeriğin zengin, hiyerarşik olarak yapılandırılmış temsillerini öğrenebileceğini kanıtladı. Sonraki yıllarda CNN mimari derinliği, eğitim verimliliği ve bilgisayarlı görü ile tıbbi görüntü yorumlaması arasındaki görev odaklı performansta üstel gelişmeler yaşandı.

Mamografide derin öğrenmenin temel avantajı, otomatik, veriye dayalı özellik öğrenme kapasitesidir. Etiketlenmiş mamografik görüntülerden oluşan geniş bir veri seti üzerinde eğitilen bir CNN, ardışık evrişim ve doğrusal olmayan aktivasyon katmanları aracılığıyla, düşük seviyeli kenarlar ve dokulardan karmaşık lezyon paternlerine ve doku bağlamına kadar giderek daha soyut temsiller çıkarmayı öğrenir. Kooi vd. (2017), 45.000'den fazla mamografik inceleme üzerinde bir CNN'yi eğiterek ve hem kitle hem de kalsifikasyon tespiti görevlerinde geleneksel CAD'den önemli ölçüde üstün performans göstererek, temel literatürle de teyit edilen büyük ölçekli çığır açıcı bir gösterim sağlamıştır.

Tıbbi görüntüleme derin öğrenme devrimini tetikleyen üç kolaylaştırıcı koşul şunlardı: büyük, küratörlü, etiketli veri setlerinin mevcudiyeti; uygun fiyatlı GPU hesaplama gücü; ve *dropout* regülarizasyonu, *batch* (yığın) normalizasyonu ve uyarlanabilir öğrenme oranı optimizasyonu dahil olmak üzere eğitim tekniklerindeki ilerlemeler. DMT yapay zekâyı anlamak için kritik bir kavram **transfer öğrenmedir (transfer learning)**: DMT inceleme veritabanları, 2B mamografi veri setlerinden önemli ölçüde daha küçük olduğundan, yalnızca DMT verileri üzerinde rastgele başlatmadan büyük derin öğrenme modellerini eğitmek aşırı uyuma (*overfitting*) eğilimlidir. Transfer öğrenimi, bol miktarda 2B veriden öğrenilen özellik temsillerini içe aktararak ve bunları hacimsel DMT alanı için rafine ederek bunu hafifletir ve böylece DMT yapay zekâ geliştirme için baskın paradigma haline gelir.

3. Mamografide Yapay Zekâ

3.1 Tek Başına (Standalone) Yapay Zekâ Performansı ve Radyologlar

Rodríguez-Ruiz vd. (2019b), 101 radyolog okuyucu ile birlikte, 653 malignite içeren 2.652 incelemeden oluşan dokuz ülkeli bir veri setinde Transpara yapay zekâ sisteminin (ScreenPoint Medical) ufuk açıcı bir değerlendirmesini gerçekleştirmiştir. Transpara sistemi 0,840'lık bir AUC'ye (%95 CI: 0,820-0,860) ulaşırken, ortalama radyolog AUC'si 0,814 (%95 CI: 0,787-0,841) olarak gerçekleşmiştir; yapay zekâ, bireysel radyologların %61,4'ünden (101 kişiden 62'si) daha iyi performans göstermiştir. Kritik bir şekilde sistem, dört farklı üreticinin dijital mamografi donanımlarında sağlam bir genelleme (*generalisation*) sergilemiş ve modern CNN'lerin donanım

değişkenliğine karşı dirençli olduğunu kanıtlamıştır. Makalenin bu çalışmayı sunumu, birincil kaynakla tam olarak örtüşmektedir.

McKinney vd. (2020), *Nature* dergisinde yayınladıkları, bir Google derin öğrenme sisteminin uluslararası değerlendirmesinde son derece bağlama bağlı olan ve bu doğrultuda yorumlanması gereken sonuçlar sunmuştur. Yalnızca tek radyolog okumasının standart bakım olduğu Amerika Birleşik Devletleri veri setinde, yapay zekâ tek okuyuculu yoruma kıyasla yanlış pozitif oranlarında mutlak %5,7'lik ve yanlış negatif oranlarında mutlak %9,4'lük önemli azalmalar sağlamıştır. Bununla birlikte, aynı yapay zekâ sistemi, iki uzman radyolog tarafından bağımsız çift okumanın zorunlu standart olduğu Birleşik Krallık veri setine uygulandığında, mutlak azalmalar yanlış pozitifler için %1,2'ye ve yanlış negatifler için %2,7'ye gerilemiştir. Bu ayrışma, klinik açıdan temel bir bulgudur: Bir yapay zekâ algoritmasının görünen etki büyüklüğü, taban çizgisi (baseline) insan iş akışının titizliği ile kesinlikle ters orantılıdır. Yalnızca tepe noktası olan ABD rakamlarının bu bağlam verilmeden raporlanması, yerel operasyonel standartlara bağlı olarak beklenen klinik yatırım getirisinin yanlış yorumlanmasına neden olabilir.

Yoon vd. (2023) tarafından yapılan ve mamografi ve DMT için tek başına yapay zekânın değerlendirildiği 16 çalışmayı kapsayan sistematik bir derleme ve meta-analiz, havuzlanmış performansın genel olarak ortalama radyolog AUC'si ile karşılaştırılabilir olduğunu bildirmiştir. Bu analizde belirlenen kritik bir uyarı, dahil edilen çalışmaların çoğunun, kanser prevalansının %10-30 ile gerçek dünyadaki %0,5-1,0 popülasyon tarama prevalansını çok aştığı, *kanser açısından zenginleştirilmiş retrospektif veri setlerini* kullanmış olmasıdır. Kanser açısından zenginleştirilmiş kohortlar üzerindeki performans, yapay zekâ doğruluğunu gerçek tarama bağlamında önemli ölçüde abartabilir ve verilerin ekstrapolasyonundan önce gerçek dünya prevalansı için kalibrasyon şarttır. Schaffter vd. (2020), yapılandırılmış bir zorluk (challenge) veri setinde birleşik yapay zekâ-radyolog değerlendirmesini incelemiş ve yapay zekâ-radyolog kombinasyonlarının her iki modalitenin bağımsız halinden daha iyi performans gösterdiğini ortaya koymuştur.

3.2 Eşzamanlı Karar Destek Sistemi Olarak Yapay Zekâ

Mamografide klinik olarak en çok değerlendirilen yapay zekâ dağıtım modeli, yapay zekâ tarafından oluşturulan işaretlerin ve şüphe skorlarının radyolog görüntü okumasıyla eşzamanlı olarak görüntülendiği "eşzamanlı karar destek"tir. Rodríguez-Ruiz vd. (2019a), 14 radyoloğun 240 mamografik incelemeyi eşzamanlı yapay zekâ desteğiyle ve desteksiz olarak yorumladığı kontrollü bir okuyucu çalışması yürütmüştür. yapay zekâ desteğiyle okuma,

ortalama okuyucu EAA'sını (AUC) 0,87'den 0,89'a ($p = 0,002$) ve duyarlılığı %83'ten %86'ya yükseltirken, okuma süresi vaka başına yaklaşık 148 saniyede sabit kalmıştır; bu da yapay zekâ desteğinin bilişsel yük eklemeyen dikkat kaynaklarını gerçekten şüpheli bulgulara yönlendirdiğini doğrulamaktadır. Samulski vd. (2010), retrospektif ikinci okuma kullanımı yerine ilk okuma sırasındaki eşzamanlı kullanımın radyolog performans faydasını en üst düzeye çıkardığını ortaya koymuştur.

Yılmaz vd. (2025), İstanbul'da yürütülen bir klinik pratikte Lunit INSIGHT MMG sistemiyle değerlendirilen 1.063 mamografik incelemeye ait prospektif, tek merkezli bir çalışma sunmuştur. Çoğu yapay zekâ validasyonunun retrospektif kalması ve retrospektif kanser zenginleştirilmiş veri setlerinin canlı klinik popülasyonlardan sistematik olarak farklılık göstermesi göz önüne alındığında, bu prospektif gerçek dünya tasarımı özellikle değerlidir. Çalışma, yapay zekânın etkili bir teşhis aracı olarak işlev gördüğünü, özellikle maskeleyen etkilerinin hem insani hem de makine algısını zorladığı yoğun meme dokusuna sahip kohortlarda radyologun tanılabilirliğini önemli ölçüde artırdığını ve okuyucular arası değişkenliği azalttığını göstermiştir. Arıbal vd. (2019), işbirlikçi yapay zekâ destekli yaklaşımların katılımcı kurumlarda erken evre meme kanseri tespitini iyileştirdiğini gösteren tamamlayıcı, çok ülkeli gözlemsel kanıtlar sağlamıştır.

3.3 İnterval Kanser Tespiti İçin yapay zekâ

Planlanmış tarama turları arasında teşhis edilen ve önceki taramada okült (gizli) olan veya daha da önemlisi görünür olduğu halde tespit edilemeyen interval kanserler, koruyucu tarama paradigmasının nihai başarısızlığını ve herhangi bir tarama programı için birincil kalite metriğini temsil eder. Bu tümörler sıklıkla agresif biyolojik fenotipler, hızlı proliferasyon ve dolayısıyla daha kötü prognoz sergilerler.

Çelik vd. (2023), Türk tarama kohortunda özellikle interval kanser tespiti için Transpara yapay zekâ sisteminin ardışık iki versiyonunu karşılaştıran retrospektif bir değerlendirme gerçekleştirmiştir. Yeni yazılım versiyonu, aynı vaka setindeki önceki versiyona göre anlamlı ölçüde iyileştirilmiş EAA (AUC) göstererek, yinelenmeli yazılım güncellemelerinin gerçek, ölçülebilir klinik kazanımlara dönüştüğünü doğrulamış ve yapay zekâ performansının art arda gelen geliştirme döngüleriyle iyileşmeye devam edeceği beklentisini valide etmiştir. Bu çalışma, spesifik olarak zorlu interval kanser alt kümesinde yapay zekâ performansını nicelendiren ilk Türk araştırmalarından birini temsil etmektedir.

Çelik ve Arıbal (2024), bu çalışmayı Türkiye'nin ulusal mamografik tarama programını kapsayacak şekilde genişletmiş, başlangıçta normal (BI-RADS 1 ve 2) olarak yorumlanan (2,1 milyondan fazla önceden negatif incelemeyi temsil eden) geniş bir mamogram kohortunu ulusal kanser kayıt sistemi ile eşleştirerek interval kanser vakalarını izole etmiştir. Transpara yapay zekâ sistemi bu kohorta uygulandığında, insanlar tarafından gözden kaçırılan interval kanserlerin %53'üne mümkün olan en yüksek şüphe skorunu (10 üzerinden 10) atamıştır. %90'lık konservatif bir özgüllük (spesifisite) eşliğinde çalışan algoritma, bu okült lezyonları tespit etmede %53,7'lik bir duyarlılığa (sensitivite) ulaşmıştır. Bu doğrulanmış bulgunun sarsıcı klinik etkileri vardır: İnsan uzmanlar tarafından görünmez veya iyi huylu kabul edilen agresif interval kanserlerin yarısından fazlası, CNN'nin başarılı bir şekilde haritalandırdığı piksel seviyesinde matematiksel malignite mimarilerine sahiptir. Yapay zekânın ileriye dönük ikincil okuyucu veya geriye dönük kalite güvence aracı olarak uygulanması, kaynakları kısıtlı ulusal programlarda interval kanser oranını yarıya indirmeye yönelik matematiksel olarak kanıtlanmış bir yol sunmaktadır.

Uluslararası kanıtlar da bu bulguları desteklemektedir. Lång vd. (2021) retrospektif bir Avrupa kohortunda yapay zekânın okuyucu radyolog tarafından negatif olarak belirlenen lezyonları inceleyerek interval kanserlerin %35'e kadarını geçmişe dönük olarak önceki tarama incelemesinde şüpheli olarak tanımlayabildiğini göstermiştir. Graewingholt ve Rossi (2021), FFDM okuma sürecine bir yapay zekâ algoritmasının eklenmesinin retrospektif bir analizde interval kanser oranını azalttığını bildirerek, yapay zekânın program kalite güvencesindeki rolünü pekiştirmiştir.

3.4 Tarama Triyajı ve İş Yüğü Azaltımı İçin yapay zekâ

Mamografide yapay zekânın operasyonel olarak en dönüştürücü uygulaması, okuma kaynaklarının verimli tahsisi için tarama muayenelerinin triyajıdır. yapay zekâ her muayeneye bir şüphe skoru atar; düşük riskli muayeneler tek okumaya yönlendirilir veya radyolog çalışma listesinden (worklist) çıkarılırken, yüksek riskli muayeneler tam çift okumaya tabi tutulur. Başlıca üç triyaj senaryosu değerlendirilmiştir: yapay zekâ, kansersiz muayeneleri insan okumasından tamamen filtreler; yapay zekâ düşük şüpheli muayeneler için ikinci insan okuyucunun yerini alır; veya yapay zekâ, muayeneleri şüphe skoruna göre çalışma listesinde yeniden sıralar.

Lång ve arkadaşları (2023) tarafından raporlanan dönüm noktası niteliğindeki Yapay Zekâ Destekli Mamografi Taraması (MASAI) çalışması, yapay zekâ destekli triyaj için en kaliteli prospektif randomize kanıtı sunmaktadır. Klinik güvenlik analizi, İsveç'teki altı tarama merkezinde randomize edilen 80.033

kadınlardan oluşan bir kohorta dayanıyordu. yapay zekâ destekli müdahale kolu (39.996 kadın) 244 kanser tespit ederken (1.000'de 6,1), standart çift okuma kontrol kolu (40.024 kadın) 203 kanser tespit etmiştir (1.000'de 5,1); bu durum yapay zekâ destekli tek okumayla taramada saptanan kanserlerde %20'lik bir artışa işaret etmektedir. Araştırmada Transpara sürüm 1.7.0 kullanılmış ve 1-10 risk skorlu triyaj mimarisi uygulanmıştır: 1'den 9'a kadar puan alan muayeneler (toplam tarama hacminin %92,0'ını oluşturur) tek okumaya yönlendirilirken, yalnızca maksimum 10 puan alanlar çift okumayı tetiklemiştir. Radyolog iş yükü tam olarak %44,3 oranında azalmıştır (kontrol kolundaki 83.231 okumaya karşılık yapay zekâ kolunda 46.345 okuma). Kritik olarak, yapay zekâ kolunda tespit edilen kanserlerin %85,2'si (244 vakadan 208'i), taranan popülasyonun yalnızca %7,2'sini oluşturan Skor 10 kategorisinde toplanmıştır. Bu derin konsantrasyon, yapay zekâ triyajının temel önermesini doğrulamaktadır: derin öğrenme modelleri, sürekli malignite olasılık dağılımlarını biyolojik gerçeklikle doğru bir şekilde eşleyebilir. Okuyucular, 105.000 kadından oluşan tam MASAI kohortundan elde edilen ve sonradan yayınlanan bir analizin %29 kanser tespiti artışı bildirdiğini not etmelidir; bu rakam atıfta bulunulan 2023 tarihli Lång vd. makalesinde yer almamaktadır ve o yayına atfedilmemelidir.

Lauritzen ve arkadaşları (2022), Danimarka'da 114.000 kadın genelinde yapay zekâ tabanlı bir mamografi tarama protokolünü raporlamış, normal incelemeleri tek okumaya ve yüksek şüpheli incelemeleri çift okumaya triyaj etmiştir. Protokol, program hassasiyetini korurken radyolog iş yükünü anlamlı ölçüde azaltmıştır. Rodríguez-Ruiz vd. (2019c), klinik olarak kabul edilemez bir hassasiyet kaybı olmadan %30-50 iş yükü azalmaları gösteren önceki bir çalışmada yapay zekâ tabanlı normal inceleme kimlik doğrulamasının fizibilitesini kanıtlamıştır.

Kızılğdağ Yirgin ve arkadaşları (2022), on yıllık Bahçeşehir Mamografik Tarama Programı bağlamında yapay zekâ tanı performansını değerlendirmiştir. Çalışma metodolojisini tam olarak not etmek esastır: Araştırmacılar yapay zekâyı tam on yıllık boylamsal tarama hacmine dağıtmamışlardır. Bunun yerine, minimum 24 aylık takibi olan 101 negatif kontrol mamogramının yanı sıra 110 teyit edilmiş maligniteyi (74 tarama tespitli, 27 interval ve 9 tarihsel olarak gözden kaçmış kanser) içeren 211 dijital mamogramdan oluşan, kanser açısından son derece seçilmiş, zenginleştirilmiş bir alt kohortu değerlendirmişlerdir. Bu vaka-kontrol tasarımı, prediktif değerleri %0,5-1,0'lık gerçek dünya tarama prevalansına kıyasla matematiksel olarak bozan %52'lik bir kanser prevalansı sağlar. Bu zenginleştirilmiş laboratuvar yapısı içinde, bağımsız yapay zekâ 0,853'lük bir AUC'ye ulaşmış ve kanserlerin %72,7'sini %34,5'lik optimal bir risk skoru eşliğinde tespit ederek, insan radyologların

%67,3'lük tespit oranını aşmıştır. Eşzamanlı üçüncü bir okuyucu olarak bir radyolog ile yapay zekâyı eşleştiren varsayımsal bir takım oluşturma modeli, toplam kanser tespitini %83,6'ya yükseltmiştir; yapay zekâ ayrıca interval kanserlerin %44,4'ünü ve geçmişte gözden kaçan kanserlerin %66,6'sını bulmuştur. Bu sonuçlar yapay zekâ etkinliğini doğrularken, zenginleştirilmiş çalışma tasarımı açık bir prevalans düzeltmesi olmadan popülasyon taramasına doğrudan uyarlanamaz.

Salim ve arkadaşları (2020), Transpara ve Lunit INSIGHT dâhil olmak üzere üç ticari yapay zekâ algoritmasının harici değerlendirmesi aracılığıyla üreticiden bağımsız karşılaştırmalı kanıtlar sunarak, ticari sistemler arasında performans farklılıkları olduğunu ve algoritma seçiminin birbirinin yerine kullanılamayacağını göstermiştir. Kim ve arkadaşları (2020), yapay zekâ kullanımının çeşitli okuyucu deneyimi seviyelerinde kanser tespit oranlarını ve yanlış pozitif geri çağırma oranlarını klinik olarak anlamlı yönlerde değiştirdiğini gösteren retrospektif, çok okuyuculu bir çalışma sunmuştur.

3.5 Meme Yoğunluğu (Densitesi) Değerlendirmesi ve Risk Tahmini

Mamografik meme yoğunluğu çift yönlü klinik önem taşır. Radyolojik olarak, yoğun fibröglandüler doku X-ışını kontrastını azaltır ("maskeleme etkisi") ve mamografik duyarlılığı yağlı memelerde %85'in üzerinden, son derece yoğun memelerde %45-65'e kadar düşürür. Epidemiyolojik olarak yoğunluk bağımsız bir risk faktörüdür: Son derece yoğun memelere (Meme Görüntüleme Raporlama ve Veri Sistemi [BI-RADS] Kategori D) sahip kadınlar, neredeyse tamamen yağlı memelere (BI-RADS Kategori A) sahip kadınlara kıyasla yaklaşık dört ila altı kat meme kanseri riski taşır. Öneme rağmen, radyologlar tarafından yapılan geleneksel BI-RADS yoğunluk sınıflandırması, muayenelerin yaklaşık %20-30'unun farklı okuyucular tarafından farklı kategorilere atanmasıyla belirgin bir okuyucular arası değişkenlik gösterir.

Yapay zekâ, standartlaştırılmış, tekrarlanabilir yoğunluk değerlendirme sunar. Dembrower ve arkadaşları (2019), tam mamografik görüntüden çıkarılan derin özellikleri içeren derin öğrenme tabanlı bir risk skorunun, yalnızca geleneksel yoğunluk kategorilerinde kodlananların ötesinde riskle ilgili bilgileri yakaladığını göstermiştir. Yala ve arkadaşları (2019), kişiselleştirilmiş tarama aralığı belirleme açısından önemli çıkarımları olan, yaygın olarak kullanılan Tyer-Cuzick klinik risk modelinden daha iyi performans gösteren ve mamografiye dayalı bir derin öğrenme risk modeli geliştirip doğrulamıştır. Ha ve arkadaşları (2019), mamografik veri setlerinden CNN tabanlı meme kanseri riski katmanlaştırması (stratifikasyonu) sergileyerek derin öğrenmenin, salt görsel incelemeyle görünmeyen riskle ilgili görüntü özelliklerini çıkardığını teyit etmiştir.

3.6 Mamografik Pozisyonlama Değerlendirmesi İçin Yapay Zekâ

Optimum mamografik pozisyonlama, tanısal görüntü kalitesi için bir ön koşuldur. Suboptimal pozisyonlama meme parankimini gizler, lezyonları taklit eden veya maskeleyen artefaktlara yol açar ve geri çağırmalara, gereksiz radyasyona ve hasta kaygısına yol açan teknik olarak kabul edilemez muayeneler üretir. Tanyel ve arkadaşları (2024), CNN'lere mutlak görüntü koordinatları farkındalığı kazandıran uzamsal derin öğrenme modelleri olan gelişmiş Koordinat Dikkat UNet (CoordAtt UNet) mimarilerini kullanarak otomatik bir pozisyonlama değerlendirme protokolü geliştirmiştir. Algoritma, mediolateral oblik (MLO) görüntülerinde meme başı ve pektoralis kasının alt marjı dahil olmak üzere kritik anatomik simge yapıları doğru bir şekilde lokalize eder, posterior meme başı çizgisini (PNL) otomatik olarak hesaplar ve nicel, açıklanabilir pozisyonlama kalite puanları oluşturur. Modalite konsoluna konuşlandırılan bu sistem, radyografe anında geometrik geri bildirim sağlayarak, hasta görüntüleme süitinden ayrılmadan önce konumlandırma hatalarının düzeltilmesini sağlar ve tüm tarama programının tanısal taban çizgisini sistematik olarak yükseltir. Brahim ve arkadaşları (2022), otomatik konumlandırma kalitesi değerlendirmesinin bir Avrupa tarama bağlamında uygulanabilirliğini göstererek, derin öğrenme tabanlı konumlandırma değerlendirmesini teknik olarak olgun bir uygulama olarak doğrulamıştır.

4. Dijital Meme Tomosentezinde (DMT) Yapay Zekâ

4.1 DMT'nin Avantajları ve İş Yükü Problemi

DMT, Z eksenindeki (derinlik) belirsizliği çözerek doku örtüşmesinin çoğunu ortadan kaldırır. Klinik faydaları iyi bir şekilde tespit edilmiştir: STORM araştırması, tek başına FFDM'ye kıyasla kombine DMT ve FFDM ile kanser tespit oranında %53'lük bir artış ve yanlış pozitif geri çağırmalarda bir azalma göstermiştir (Ciatto vd., 2013). Haas ve arkadaşları (2013), bir Birleşik Devletler toplum sağlığı uygulaması ortamında artan kanser tespiti ve azalan geri çağırma doğrularken; Friedewald ve arkadaşları (2014) 450.000'den fazla tarama incelemesini analiz ederek, kombine DMT ve FFDM okumasıyla invaziv kanser tespitinde %41'lik bir artış ve geri çağırma oranında %15'lik bir azalma olduğunu kanıtlamıştır.

Bu faydaların bir operasyonel bedeli vardır. Tagliafico ve arkadaşları (2017), altı okuma stratejisi genelinde DMT için okuma sürelerini ölçmüş ve FFDM'nin %50-200 üzerinde sürekli olarak daha uzun yorumlama süreleri göstermiştir; uzman okuyucular daha az deneyimli okuyuculardan orantılı olarak daha yüksek verimlilik elde edememiştir. Yüksek hacimli tarama ortamlarında, okuma süresindeki bu artış doğrudan okuyucu veriminin azalmasına, muayene

başına maliyetin artmasına veya mevcut radyolog kapasitesini aşan iş gücü gereksinimlerine dönüşür.

Üretici (vendor) heterojenliği, DMT'ye özgü ek bir teknik zorluk ortaya koyar. Normal meme parankiminin DMT görüntülerindeki görünümü, açılal aralık, dilim kalınlığı, rekonstrüksiyon algoritmaları ve gürültü özelliklerindeki farklılıklar nedeniyle Hologic, GE HealthCare, Siemens Healthineers ve Fujifilm gibi üreticiler arasında önemli ölçüde farklılık gösterir. DMT'deki bu üreticiler arası görünüm değışkenliği, FFDM sistemlerindeki ilgili değışkenlikten önemli ölçüde daha büyüktür ve gerçek dünyadaki ekipman karışmaları genelinde genelleştirilebilmesi gereken yapay zekâ modellerini eğitmek için ciddi bir engel yaratır. Ağırlıklı olarak tek bir üreticinin DMT görüntüleri üzerinde eğitilmiş bir yapay zekâ modeli, rakip üreticinin sisteminde optimal olmayan performans gösterebilir. Üreticiler arası (cross-vendor) validasyon, ticari yapay zekâ validasyon çalışmalarının temel ancak sıklıkla atlanan bir bileşenidir.

4.2 DMT İçin Yapay Zekâ-CAD Sistemleri: Teknoloji ve Doğrulanmış Performans

Erken dönem DMT yapay zekâ sistemleri, transfer öğrenme yoluyla 2B mamografi ağlarının doğrudan uyarlamalarıydı. Modern derin öğrenme sistemleri, dilimler arasındaki uzamsal sürekliliğin açıkça modellenmesine olanak tanıyarak tam DMT hacmini üç boyutlu veya yarı 3B evrimsel mimariler aracılığıyla işler. Mevcut nesil sistemlerden alınan klinik çıktılar arasında rekonstrükte dilimler üzerine yerleştirilmiş dilim başına lezyon ısı haritaları, sürekli bir ölçekte muayene düzeyinde malignite skorları ve sentetik 2B görüntüler üzerindeki algılama işaretleri bulunur. FDA, Mart 2020'de DMT yorumlaması için Transpara'ya (ScreenPoint Medical) izin vermiştir.

Park ve arkadaşları (2024), ABD ve Güney Kore'deki 14 kurumdan retrospektif olarak toplanan verilerle DMT için bir derin öğrenme yapay zekâ algoritması geliştirmiş ve çok merkezli, çok okuyuculu bir çalışmada değerlendirmiştir. DMT üzerinde değerlendirilen bağımsız yapay zekâ algoritması, 0,93 (%95 CI: 0,92-0,94) gibi dikkate değer bir EAA'ya (AUC) ulaşmıştır. Çok okuyuculu değerlendirme bileşeninde, yapay zekâ, radyologların %77,34'lük oranına kıyasla %89,64'lük üstün bir özgüllük (spesifisite) göstermiştir ($p < 0,001$). Eşzamanlı karar desteği olarak entegre edildiğinde yapay zekâ, insan okuyucuların EAA'sını 0,90'dan 0,92'ye ($p = 0,003$) ve duyarlılığı %85,4'ten %87,7'ye çıkarmıştır. Bu birincil veriler, karmaşık derin öğrenme mimarilerinin DMT görüntü yığınlarının karmaşık uzamsal geometrilerinde başarıyla ustalaştığını doğrulamaktadır.

Mevcut DMT yapay zekâ sistemlerinin önemli bir teknik kısıtlaması, mikro kalsifikasyon tespiti için optimal olmayan performanslarıdır. Görüntü alımının sınırlı açılmalı aralığı nedeniyle, DMT'nin Z eksenli uzamsal çözünürlüğü (derinlik boyutundaki çözünürlük), düzlem içi (XY) çözünürlüğünden önemli ölçüde daha düşüktür. 0,1 mm kadar küçük olabilen ve üç boyutlu uzamsal dağılımları malignite için önemli bir tanı kriteri olan malign mikro kalsifikasyonlar, birden fazla dilime bulaşabilir veya rekonstrüksiyon artefaktları tarafından gizlenerek dilim başına derin öğrenme tespitini güvenilir hale getirebilir. Mikro kalsifikasyonlar duktal karsinoma in situ (DCIS) için birincil gösterge olduğundan, DMT yapay zekâdaki kalsifikasyon tespiti açığı, bu sistemleri kullanan radyologlara açıkça iletilmesi gereken klinik olarak anlamlı bir kısıtlamayı temsil eder.

4.3 DMT'de Yapay Zekâ Performansı — Temel Çalışmalar

Van Winkel vd. (2021), değişen deneyime sahip okuyucular ve 240 muayene üzerinden DMT yorumlamasında Transpara yapay zekâ desteğinin etkisini değerlendiren çok okuyuculu, çok vakalı bir çalışma yürütmüştür. Yapay zekâ desteği, hem acemi hem de deneyimli radyologlar için ortalama okuyucu EAA'sını iyileştirmiş ve okuma süresini önemli ölçüde azaltmıştır. Okuma süresindeki azalmanın 2B eşzamanlı destek literatüründe gösterilmemiş olması önemli bir gözlemdir, zira bu durum DMT'nin hacimsel karmaşıklığının özellikle dilim yığını (slice stack) boyunca yapay zekâ yönlendirmeli dikkatsel odaklanmadan fayda sağladığını düşündürmektedir.

Conant ve arkadaşları (2019), eşzamanlı yapay zekâ kullanımı ile doğruluğun ve verimliliğin arttığını bildirerek DMT yorumlamasında spesifik yapay zekâ faydasının ilk prospektif demonstrasyonlarından birini sağlamıştır. Li ve arkadaşları (2020), DMT ve FFDM girdilerini birleştiren derin öğrenme sınıflandırmasının her iki modaliteden tek başına daha iyi performans gösterdiğini kanıtlamıştır –bu, hem DMT hem de geleneksel 2B verilerinin eşzamanlı olarak mevcut olduğu entegre okuma iş akışları için önemli bir bulgudur. yapay zekâ ile geliştirilmiş DMT mamografisi: Transpara 1.7.0 ve ProFound AI 3.0, AUC ve yoğunluk alt grup performansını değerlendiren retrospektif çalışmalarda DMT için karşılaştırılmış; algoritma seçiminin klinik sonuçları etkilediği ve üreticiden bağımsız (vendor-neutral) kanıtların gerekli olduğu pekiştirilmiştir.

4.4 DMT Taramasında İş Yükü Azaltımı İçin Yapay Zekâ

Raya-Povedano ve arkadaşları (2021), bir İspanyol toplum tabanlı programında 15.987 eşleştirilmiş dijital mamografi ve DMT muayenesinin retrospektif bir değerlendirmesi olan Córdoba Tomosentez Tarama Çalışmasını

raporlamıştır. yapay zekâ, %5'i aşan bir duyarlılık düşüşü olmaksızın DMT muayenelerinin %70'ine kadar olan kısım için okuma gereksinimini güvenli bir şekilde ortadan kaldırmıştır. En klinik olarak ilgili karşılaştırma, geleneksel 2B mamografi çift okumasından birleşik yapay zekâ artı DMT okumasına geçişin, eşzamanlı olarak %25 duyarlılık artışı, %27 geri çağırma oranı azalması ve %30 toplam iş yükü azalması sağladığını göstermiştir; bu da yapay zekâ triyajının toplum düzeyinde DMT taramasını uygulanabilir kılmak için gereken operasyonel köprü olduğunu doğrulamıştır.

Dahlblom ve arkadaşları (2023) İsveç'teki bir tarama çalışmasından 14.772 çift okumalı DMT muayenesini analiz etmiştir. Yalnızca yapay zekâ tarafından şüpheli olarak işaretlenen muayenelerin %50,5'ini okumak, tam çift okuma ile tespit edilen kanserlerin %95'ini yakalamış ve %40-70 iş yükü azaltımlarıyla yapay zekânın ikinci okuyucunun yerini almasını operasyonel olarak en uygun strateji olarak tanımlamıştır. Romero-Martín ve arkadaşları (2022), tek başına çalışan yapay zekânın bir İspanyol tarama programında hem FFDM hem de DMT için ikinci okuyucu duyarlılığına uyduğunu veya bunu aştığını göstermiştir. Magni ve arkadaşları (2023), DMT için yapay zekânın kapsamlı bir incelemesini sunarak iş yükü azalmasının bir dizi program yapısı genelinde elde edilebilir olduğunu doğrulamıştır. Dahlblom ve arkadaşları (2023), iş yükü azaltımı ve saptama çıktılarının belirli okuma protokolü tasarımı seçimlerine olan duyarlılığını kanıtlamıştır.

4.5 DMT Yapay Zekâya Özgü Teknik Zorluklar

Beş teknik açıdan belirgin zorluk, mevcut DMT yapay zekâ sistemlerinin performansını sınırlar ve uygun klinik uygulamanın sağlanması için meme radyologları tarafından anlaşılması gerekir.

Birincisi, Z eksenli çözünürlük düşüklüğüdür. DMT rekonstrüksiyonları, düzlem içi 0,05-0,10 mm piksel boyutuna kıyasla, tipik olarak dilim merkezleri arasında 1 mm olacak şekilde derinlik çözünürlüğünde önemli ölçüde daha zayıftır. Bireysel dilimleri bağımsız olarak işleyen yapay zekâ sistemleri, lezyon morfolojisinin üç boyutlu bağlamını yakalayamaz ve tüm yığını işleyenler hacimsel veri boyutunun getirdiği hesaplama kısıtlamalarıyla karşılaşır. Dilimler arasında bilginin entegrasyonu, hem kitle hem de kalsifikasyon tespiti doğruluğunu etkileyen çözülmemiş bir mimari zorluk olmaya devam etmektedir.

İkincisi, üretici heterojenliğidir. Bölüm 4.1'de belirtildiği gibi, DMT görünümü üreticiler arasında FFDM üreticilerine kıyasla çok daha fazla değişir. Ağırlıklı olarak bir üreticinin verileri üzerinde eğitilmiş bir model, diğer üreticilerin sistemlerinde düşük performans gösterir. Klinik olarak uygulanabilir

yapay zekâ, hedef tarama ortamındaki tam ekipman karışımını temsil eden çok üreticili veri setlerinde eğitilmeli ve onaylanmalıdır. Mevcut ticari yapay zekâ doğrulamaları tek tip olarak bu standardı karşılamaz ve yapay zekâyı eğitim verilerinde temsil edilmeyen ekipmanlarda dağıtan kurumlar için bir uyarı oluşturur.

Üçüncüsü, mikro kalsifikasyon tespit açığıdır. DMT yapay zekâda mikro kalsifikasyon tespiti, Z eksenli çözünürlük kısıtlaması ve kalsifikasyonların rekonstrükte edilmiş dilim kalınlığına kıyasla küçük boyutu nedeniyle kitle ve distorsiyon tespiti kadar güvenilir değildir. Klinik uygulamada radyologların, özellikle kalsifikasyon değerlendirmesi için DMT yapay zekâyı geleneksel 2B FFDM okumasıyla desteklemesi gerekebilir; bu, program protokolü tasarımında ileriye dönük olarak ele alınması gereken bir iş akışı komplikasyonudur.

Dördüncü zorluk, zamansal boşluktur — önceki görüntü karşılaştırmasının olmaması. Bir meme radyoloğu bir muayeneyi yorumlarken, mevcut görüntüleri rutin olarak önceki muayenelerle karşılaştırır ve mikroskobik yapısal değişiklikleri, bir kitlenin yavaş asimetric büyümesini veya ilerleyen parankimal arka plan değişikliklerini saptamak için etkili bir şekilde görsel çıkarma gerçekleştirir. Mevcut ticari yapay zekâ algoritmaları katı bir zamansal boşlukta çalışır: Mevcut muayenenin piksel verilerini tamamen kesitsel olarak, boylamsal bağlam olmadan çözümlerler. Sonuç olarak yapay zekâ, tek bir statik görüntüde matematiksel olarak iyi huylu kaydedilen yavaş gelişen maligniteleri saptamada insan uzmanlara kıyasla doğası gereği düşük performans gösterir. Rijit olmayan boylamsal görüntü kaydı ve piksel düzeyinde zamansal çıkarma yeteneğine sahip uzamsal-zamansal CNN mimarileri ticarileşene kadar, boylamsal değişimi izlemede insan okuyucunun teşhis üstünlüğü tehdit edilmemektedir.

Beşinci zorluk, sentetik 2B entegrasyonudur. DMT verilerinden (C-view, SmartSlices) üretilen sentetik 2B mamogramlar, alınan FFDM'nin yanında veya yerine giderek daha fazla kullanılarak radyasyon dozunu düşürmektedir. Yerel (native) FFDM'de onaylanan yapay zekâ modelleri, gürültü özellikleri, uzamsal frekans içeriği ve doku görünümü bakımından farklılık gösteren DMT'den üretilen sentetik 2B görüntülerde eşdeğer performans göstermeyebilir. FFDM ile doğrulanmış yapay zekâyı sentetik 2B iş akışlarında dağıtmadan önce sentetik 2B görüntüler üzerinde özel bir doğrulama yapılması gerekir.

4.6 Multimodalite Sinerjisi: Otomatik Meme Ultrasonu ile Maskeleme Etkisinin Aşılması

Yoğun fibrögländüler doku, X-ışını penetransı için aşılabilir bir fiziksel engel oluşturarak hem insan okuyucuların hem de mamografi ve DMT'deki

yapay zekâ algoritmalarının duyarlılığını azaltır. Yoğun memelere sahip kadın alt kümesinde bu maskeleye etkisini atlatmak için tamamlayıcı tarama modalitelerine ihtiyaç vardır.

Arıbal vd. (2024), DMT'ye göre bağımsız ve tamamlayıcı bir tarama modalitesi olarak otomatik meme ultrasonunun (ABUS) etkinliğini değerlendirmiştir. Doğrulanmış veriler, tek başına ABUS'un yaklaşık 0,88'e yaklaşan bir havuzlanmış duyarlılıkla DMT'ye (1.000 taramada 7,5) kıyaslanabilir kanser tespit oranlarına ulaştığını doğruladı. DMT üstün bir özgüllük profiline (spesifisite) sahip olmaya devam etse de (%95'e karşı %88), ABUS yoğun dokunun gizlediği yüksek derecede invaziv kanserleri tespit etmede benzersiz bir yeterlilik gösterdi. Teşhis amaçlı yapay zekânın mantıksal bir sonraki evrimi çok modlu (multi-modal) füzyon mimarilerinde yatmaktadır — DMT'nin hacimsel uzamsal verilerini ABUS'un mekanik akustik doku özellikleriyle senkronize olarak ayrıştırabilen sistemler; böylece yoğun memelerde X-ışını modalitelerinin yüksek özgüllüğü ile ultrasonun yüksek duyarlılığını sentezler. Bu çok modlu yaklaşım, nihayetinde son derece yoğun memelere sahip kadınlar için en uygun tarama yolunu (pathway) tanımlayabilir.

5. Klinik Uygulama ve Kanıtlar

5.1 Araştırma Örnekleri

Mamografi ve DMT'de yapay zekâ için kanıt temeli, küçük, tek kurum tabanlı, kanser açısından zenginleştirilmiş retrospektif çalışmalardan; daha büyük, çok merkezli, prospektif ve randomize değerlendirmelere doğru ilerleyen bir evrim ile karakterizedir. %10-30 oranındaki kanser prevalansının gerçek dünyadaki %0,5-1,0 oranındaki popülasyon taraması prevalansını fazlasıyla aştığı, kanserden zenginleştirilmiş retrospektif çalışma tasarımlarının temel kısıtlılığı; bu veri setlerinde kalibre edilen yapay zekâ sistemlerinin gerçek tarama bağlamında eşdeğer performans gösteremeyebileceğidir. Kanser prevalansı düşük olduğunda, özgüllükteki (spesifisite) küçük değişiklikler bile çok sayıda yanlış pozitif geri çağırmaya neden olur ve kanser açısından zenginleştirilmiş kohortlardaki performans, prevalans düzeltmesi olmadan doğrudan ekstrapole edilemez.

MASAI çalışması (Lång vd., 2023), prospektif randomize kanıtlara geçişte kritik kilometre taşını temsil etmektedir. Eşzamanlı iş yükü azalmasıyla (%44,3) birlikte kanser tespiti gelişiminin boyutu (2023 güvenlik kohortunda çift okumanın %20 üzerinde), prospektif tarama koşullarında yapay zekâ faydasının klinik olarak anlamlı olduğunu göstermektedir. Álvarez-Benito vd. (2023), yapay zekânın gerçek klinik pratikte hem 2B mamografi hem de DMT için eşzamanlı destek olarak uygulandığı İspanyol toplum tabanlı bir tarama

programından ileriye dönük kanıtlar sunmuş, yapay zekânın her iki modalitede de eşzamanlı olarak kanser tespit oranını ve geri çağırımların pozitif prediktif değerini iyileştirdiğini göstermiştir.

5.2 Kullanım Entegrasyon Modelleri

Meme taramasında yapay zekânın klinik uygulaması, her birinin farklı kanıt temelleri, iş yükü etkileri ve risk profilleri olan üç ana operasyonel modele göre yapılandırılabilir.

Lauritzen vd. (2022) ve Rodríguez-Ruiz vd. (2019c) kaynaklı kanıtlarla desteklenen triyaj/filtreleme modeli, mutlak iş yükü azaltımı açısından en büyük potansiyeli sunar ve ciddi radyolog iş gücü kısıtlamalarıyla karşı karşıya olan programlar için en uygun olanıdır. Temel kısıtlılığı, yapay zekânın gerçek bir kansere az puan (under-score) verme ihtimali ile çift yorumlama gerektiren bir inceleme için tek okuma yapılmasına yol açma riskidir. Titiz güvenlik eşiği kalibrasyonu ve yapay zekâ-düşük-riskli kaçırılan kanser oranlarının sürekli izlenmesi temel program yönetişimi gereksinimleridir.

Rodríguez-Ruiz vd. (2019a) ve Van Winkel vd. (2021) tarafından desteklenen eşzamanlı karar destek modeli, en elverişli risk profiline sahip en kanıta dayalı kısa vadeli uygulama yolunu sunar: Radyologlar tam yorumlayıcı özerkliklerini korurlar ve yapay zekâ yanlış sınıflandırması temel iş akışını değiştirmez. İkinci okuyucu değişimi modeli (Dahlblom vd., 2023; Lång vd., 2023), yerleşik çift okuma altyapısına sahip programlar için en operasyonel dengeyi sunar ancak çoğu yargı bölgesinde yasal çerçeveler çözülmemiş durumdadır.

Tüm modellerde kritik bir yönetim sorunu, radyologların yorumlama yargısını yapay zekâya aşırı devretme ve yeterli bağımsız eleştirel değerlendirme yapmaksızın yapay zekâ kararlarını kabul etme eğilimi olan otomasyon önyargısıdır. Radyologların insan üstü yetenekleri zamanla zayıflayabilir ve bu, conventional QA metriklerinde görünmez olabilir. Yönetişim çerçeveleri, düzenli ‘yapay zekâ kapalı’ okuma oturumlarını ve yapay zekâ çıktısının eleştirel olarak değerlendirilmesi konusunda açık bir eğitimi zorunlu kılmalıdır.

5.3 Orta Gelirli Ülke Perspektifi

yapay zekâ doğrulama kanıtlarının ezici çoğunluğu yüksek gelirli, yüksek hacimli Avrupa ve Kuzey Amerika tarama programlarında üretilmiştir. Bu kanıtların orta gelirli ülke bağlamlarına –nüfus düzeyindeki meme yoğunluğu dağılımlarının, VKİ (BMI) profillerinin, kanser alt tipi frekanslarının, mamografik ekipman kombinasyonunun ve program altyapısının farklı olduğu ortamlara– uygulanabilirliği spesifik validasyon olmadan varsayılmaz.

Türkiye bu manzarada son derece bilgilendirici bir konuma sahiptir. Arıbal ve grubu tarafından 2009 yılında kurulan Bahçeşehir Mamografik Tarama Programı, Türkiye'nin ilk organize, toplum tabanlı mamografik tarama programıydı ve 2019 yılına kadar kesintisiz faaliyet göstermiştir. Özkan Gürdal vd. (2021), programın uzun vadeli etkililiğini ulusal kayıt verileriyle karşılaştırarak programın uluslararası ölçütlerle tutarlı bir mortalite düşüşü sağladığını doğrulamış, Avrupa taramalarının hakiki bir analogu olarak tescillenmiştir. Kızıldağ Yirgin vd. (2022), yapay zekâ tanılama performansını değerlendirmek için bu programdan derlenmiş alt kohortu kullanmış ve Batı Avrupa dışındaki ortamlarda yapay zekânın faydasına dair gerçek dünya kanıtları sağlamıştır.

Çelik ve Arıbal (2024), interval kanser tespiti için yapay zekâ performansının Türkiye'nin ulusal mamografik tarama programını kapsadığını göstermiştir —bu, radyolog kapasitesinin Batı Avrupa'daki benzerlerine kıyasla daha kısıtlı olduğu bir ülkede program kalite güvence stratejisi için acil sonuçlar doğuran bir bulgudur. Yapay zekâ triyajına karşı olan argümanın iki deneyimli radyolog tarafından çift okuma değil, kaynağı yetersiz tek radyolog tarafından tek okuma olduğu ortamlarda, yapay zekâ dağıtımının risk-fayda hesabı önemli ölçüde benimseme yönüne kaymaktadır.

6. Etik, Yasal ve Yönetişim Boyutları

Meme görüntüleme için yapay zekâ için yasal yol, FDA'nın Tıbbi Cihaz Olarak Yazılım (SaMD) çerçevesi tarafından yönetilmektedir. Derin öğrenme sistemleri karmaşık bir yasal zorluk teşkil eder; ek veriler biriktikçe yeniden eğitilebilir veya güncellenebilirler ve başlangıçtaki onaydan sonra performans özelliklerini değiştirebilirler.

Algoritmik önyargı, belirgin bir yönetim sorunu teşkil etmektedir. Eğer yapay zekâ eğitim verisetleri (yaş, ırk, vücut kitle indeksi, meme yoğunluğu vb. açısından) tarama popülasyonunun tam çeşitliliğini temsil etmiyorsa, model yetersiz temsil edilen alt gruplar için sistematik olarak düşük performans gösterecektir.

Otomasyon önyargısının ise daha derin bir etik boyutu vardır: Radyologlar okuma otoritesini yapay zekâya fazla devrederlerse, nihai güvenlik kontrolü olan insani bilişsel kapasitede sistematik bir bozulma meydana gelecektir. Bunu engellemek için kurumsal denetim mekanizmalarında radyologların içsel algı keskinliğini aktif tutmak adına kör 'yapay zekâ-kapalı' vardiyaları olmalıdır.

Sorumluluk da kritik bir husustur. Bir yapay zekâ sistemi bir kanseri işaretleyemediğinde — veya bir radyolog yapay zekâ-negatif değerlendirmesini bağımsız doğrulamadan kabul ettiğinde — yargı çerçevesi sorun yaşayacaktır.

Yönetişim ilkesi net olmalıdır: Kurallar, yapısal gözetimin önleyebileceği zararlara sonradan yanıt vermek yerine kapsamlı dağıtımdan önce belirlenmelidir.

7. Gelecek Yönelimler ve Radyoloğun Rolü

Meme görüntülemesinde yapay zekâ için kısa vadeli ufuk, klinik entegrasyon için konumlandırılmış, teknik olarak olgunlaşmış uygulamaları içerir. Pozisyonlama ve hareket artefaktlarına ilişkin gerçek zamanlı kalite tespiti, teknik sorunları henüz radyoloğun önüne gelmeden radyografe iletacaktır.

Orta vadede, multimodal yapay zekâ gelişimi ufuktaki en dönüştürücü klinik husustur. Mamografik bilgiyi klinik risk faktörleri (yaş, aile öyküsü vb.), volumetrik meme yoğunluğu ve genomik biyogöstergelerle entegre eden çok modlu bir yapay zekâ; tek kalıp protokollere yerine, gerçekten kişiselleştirilmiş bir tarama aralığı sunabilir (Yala vd., 2019; Dembrower vd., 2019).

Uzun vadede ise, “önceki görüntülerle anlamlı karşılaştırma yapabilen” sistemlerin geliştirilmesi hedeflenmektedir. Bu engeller aşıldığında meme radyoloğunun rolü “ortadan kalkmayacak”, aksine dönüşüme uğrayacaktır. Geleceğin meme radyoloğu, yapay zekânın güçlendirildiği bir ortamda yapay zekâ hatalarını ve kör noktalarını anlayan ve değerlendiren asıl “klinik hakem” olarak görev yapacaktır. Bu da mevcut asistanlık müfredatına standart bir şekilde entegre edilmesi gereken yeni bir “yapay zekâ Okuryazarlığı” gerektirir.

Temel Çıkarımlar

- **derin öğrenme tabanlı yapay zekâ, birinci nesil CAD’den temel bir geçiştir.** Büyük ölçekli doğrulamalar, mamografi ve DMT’de (0,84-0,93 AUC ile) ortalama radyolog performansını aştığını doğrulamaktadır.
- **Eşzamanlı karar desteği, en fazla kanıtla desteklenen kısa vadeli modeldir.** Radyolog AUC’sini artırırken (0,87 → 0,89) okuma süresini cezalandırmaz.
- **DMT’nin iş yükü krizi, yapay zekânın en iyi çözeceği temel sorundur.** MASAI araştırması (Lång vd., 2023), yapay zekâ tek okumanın çift okumadan %20 daha fazla kanser tespit edip, iş yükünü %44,3 azalttığını kanıtlamıştır.
- **İnterval kanser tespiti yapay zekâ için niceliksel bir etki yolu sunar.** Çelik ve Arıbal (2024), gözden kaçan interval kanserlerin %53’üne yapay zekânın en yüksek puanı (10/10) verdiğini göstererek ulusal programlarda bu oranı yarıya indirme potansiyelini kanıtlamıştır.
- **Türk popülasyon çalışmaları kanıt temelini orta gelirli ortamlara taşır.** Bahçeşehir ve ulusal tarama programları, Batı Avrupa dışındaki

popülasyonlar için operasyonel faydanın yüksek gelirli ülkelere kıyasla orantısız olarak daha da fazla olduğunu ortaya koymaktadır.

- **Doğrulanmış teknik kısıtlamalar açıkça iletilmelidir.** Üretici heterojenliği, DMT'deki mikro kalsifikasyon açığı, Z eksenini çözünürlük sorunları, zamansal boşluk (eski filmin olmaması) ve kanserden zenginleştirilmiş set enflasyonu (%52 prevalans kullanan çalışmalar) dikkatle bağlama oturtulmalıdır.
- **McKinney vd. (2020) faydası katı bir şekilde bağlama dayalıdır.** ABD (tek okuyucu) faydası yüksektir, ancak Birleşik Krallık (çift okuyucu) modelinde mutlak fayda çok düşüktür; yapay zekânın görünür faydası insan iş akışının katılığıyla ters orantılıdır.
- **Yönetişim çerçeveleri dağıtımdan önce gelmelidir.** Algoritmik önyargı ve hukuki belirsizlikler yasal düzenlemelerle güvence altına alınmalıdır.

Kaynaklar

1. Álvarez-Benito, M., Elías-Cabot, E., & Romero-Martín, S. (2023). Impact of real-life use of artificial intelligence as support for human double reading in a population-based breast cancer screening program with mammography and tomosynthesis. *European Radiology*, 34(4), 2280–2291.
2. Aribal, M. E., Mora, P., Chaturvedi, A. K., et al. (2019). Improvement of early detection of breast cancer through collaborative multi-country efforts: Observational clinical study. *European Journal of Radiology*, 115, 31–38.
3. Aribal, M. E., Seker, M. E., Guldogan, N., & Yilmaz, E. (2024). Value of automated breast ultrasound in screening: Standalone and as a supplemental to digital breast tomosynthesis. *International Journal of Cancer*, 155(8).
4. Bird, R. E., Wallace, T. W., & Yankaskas, B. C. (1992). Analysis of cancers missed at screening mammography. *Radiology*, 184(3), 613–617.
5. Brahim, M., Westerkamp, K., Hempel, L., et al. (2022). Automated assessment of breast positioning quality in screening mammography. *Cancers*, 14(19), 4704.
6. Ciatto, S., Houssami, N., Bernardi, D., et al. (2013). Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): A prospective comparison study. *The Lancet Oncology*, 14(7), 583–589.
7. Conant, E. F., Toledano, A. Y., Periaswamy, S., et al. (2019). Improving accuracy and efficiency with concurrent use of artificial intelligence for digital breast tomosynthesis. *Radiology: Artificial Intelligence*, 1(4), e180096.
8. Çelik, L., & Aribal, M. E. (2024). The efficacy of artificial intelligence (AI) in detecting interval cancers in the national screening program of a middle-income country. *Clinical Radiology*, 79(10), e885–e891.
9. Çelik, L., & Özçağlayan, Ö. (2023). Mamografi ve tomosentezde yapay zekâ. In L. Çelik & Ö. Özçağlayan (Eds.), *Meme radyolojisi* (pp. 23–26). Türkiye Klinikleri.
10. Çelik, L., Güner, D. C., Özçağlayan, Ö., Çubuk, R., & Aribal, M. E. (2023). Diagnostic performance of two versions of an artificial intelligence system in interval breast cancer detection. *Acta Radiologica*, 64(11), 2891–2897.
11. Dahlblom, V., Dustler, M., Tingberg, A., & Zackrisson, S. (2023). Breast cancer screening with digital breast tomosynthesis: Comparison of different reading strategies implementing artificial intelligence. *European Radiology*, 33(5), 3754–3765.
12. Dembrower, K., Liu, Y., Azizpour, H., et al. (2019). Comparison of a deep learning risk score and standard mammographic density score for breast cancer risk prediction. *Radiology*, 294(2), 265–272.

13. Duffy, S. W., Tabár, L., Yen, A. M.-F., et al. (2021). Beneficial effect of consecutive screening mammography examinations on mortality from breast cancer: A prospective study. *Radiology*, 299(3), 541–547.
14. Fenton, J. J., Taplin, S. H., Carney, P. A., et al. (2007). Influence of computer-aided detection on performance of screening mammography. *New England Journal of Medicine*, 356(14), 1399–1409.
15. Friedewald, S. M., Rafferty, E. A., Rose, S. L., et al. (2014). Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA*, 311(24), 2499–2507.
16. Geras, K. J., Mann, R. M., & Moy, L. (2019). Artificial intelligence for mammography and digital breast tomosynthesis: Current concepts and future perspectives. *Radiology*, 293(2), 246–259.
17. Graewingholt, A., & Rossi, P. G. (2021). Retrospective analysis of the effect on interval cancer rate of adding an artificial intelligence algorithm to the reading process for two-dimensional full-field digital mammography. *Journal of Medical Screening*, 28(4), 369–371.
18. Ha, R., Chang, P., Karcich, J., et al. (2019). Convolutional neural network based breast cancer risk stratification using a mammographic dataset. *Academic Radiology*, 26(4), 544–549.
19. Haas, B. M., Kalra, V., Geisel, J., Raghu, M., Durand, M., & Philpotts, L. E. (2013). Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening. *Radiology*, 269(3), 694–700.
20. Hickman, S. E., Baxter, G. C., & Gilbert, F. J. (2021). Adoption of artificial intelligence in breast imaging: Evaluation, ethical constraints and limitations. *British Journal of Cancer*, 125(1), 15–22.
21. Hickman, S. E., Woitek, R., Le, E. P. V., et al. (2022). Machine learning for workflow applications in screening mammography: Systematic review and meta-analysis. *Radiology*, 302(1), 88–104.
22. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, J. E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, 61(2), 69–90.
23. Kim, H. E., Kim, H. H., Han, B. K., et al. (2020). Changes in cancer detection and false-positive recall in mammography using artificial intelligence: A retrospective, multireader study. *Lancet Digital Health*, 2(3), e138–e148.
24. Kizildag Yirgin, I., Koyluoglu, Y. O., Seker, M. E., Ozkan Gurdal, S., Ozaydin, A. N., Ozcinar, B., Cabioglu, N., Ozmen, V., & Aribal, M. E. (2022). Diagnostic performance of AI for cancers registered in a mammography screening program: A retrospective analysis. *Technology in Cancer Research & Treatment*, 21.

25. Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., et al. (2017). Large scale deep learning for computer aided detection of mammographic lesions. *Medical Image Analysis*, 35, 303–312.
26. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
27. Lång, K., Hofvind, S., Rodríguez-Ruiz, A., & Andersson, I. (2021). Can artificial intelligence reduce the interval cancer rate in mammography screening? *European Radiology*, 31(8), 5940–5947.
28. Lång, K., Josefsson, V., Larsson, A. M., et al. (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. *Lancet Oncology*, 24(8), 936–944.
29. Lauritzen, A. D., Rodríguez-Ruiz, A., von Euler-Chelpin, M. C., et al. (2022). An artificial intelligence-based mammography screening protocol for breast cancer: Outcome and radiologist workload. *Radiology*, 304(1), 41–49.
30. Lehman, C. D., Wellman, R. D., Buist, D. S., Kerlikowske, K., Tosteson, A. N. A., & Miglioretti, D. (2015). Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Internal Medicine*, 175(11), 1828–1837.
31. Li, X., Qin, G., He, Q., et al. (2020). Digital breast tomosynthesis versus digital mammography: Integration of image modalities enhances deep learning-based breast mass classification. *European Radiology*, 30(2), 778–788.
32. Magni, V., Cozzi, A., Schiaffino, S., Colarieti, A., & Sardanelli, F. (2023). Artificial intelligence for digital breast tomosynthesis: Impact on diagnostic performance, reading times, and workload in the era of personalized screening. *European Journal of Radiology*, 158, 110631.
33. McKinney, S. M., Siemiek, M., Godbole, V., et al. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89–94.
34. Meyers, P. H., Nice, C. M., Becker, H. C., Nettleton, W. J., Sweeney, J. W., & Meckstroth, G. R. (1964). Automated computer analysis of radiographic images. *Radiology*, 83, 1029–1034.
35. Ozkan Gurdal, S., Ozaydin, A. N., Aribal, M. E., et al. (2021). Bahcesehir long-term population-based screening compared to National Breast Cancer Registry Data: Effectiveness of screening in an emerging country. *Diagnostic and Interventional Radiology*, 27, 157–163.
36. Park, E. K., Kwak, S., Lee, W., Choi, J. S., Kooi, T., & Kim, E. K. (2024). Impact of AI for digital breast tomosynthesis on breast cancer detection and interpretation time. *Radiology: Artificial Intelligence*, 6(3), e230318.

37. Raya-Povedano, J. L., Romero-Martín, S., Elías-Cabot, E., Gubern-Mérida, A., Rodríguez-Ruiz, A., & Álvarez-Benito, M. (2021). AI-based strategies to reduce workload in breast cancer screening with mammography and tomosynthesis: A retrospective evaluation. *Radiology*, 300(1), 57–65.
38. Rodríguez-Ruiz, A., Krupinski, E., Mordang, J. J., Schilling, K., Heywang-Köb-runner, S. H., Sechopoulos, I., & Mann, R. M. (2019a). Detection of breast cancer with mammography: Effect of an artificial intelligence support system. *Radiology*, 290(2), 305–314.
39. Rodríguez-Ruiz, A., Lång, K., Gubern-Mérida, A., et al. (2019b). Stand-alone artificial intelligence for breast cancer detection in mammography: Comparison with 101 radiologists. *Journal of the National Cancer Institute*, 111(9), 916–922.
40. Rodríguez-Ruiz, A., Lång, K., Gubern-Mérida, A., et al. (2019c). Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. *European Radiology*, 29(9), 4825–4832.
41. Romero-Martín, S., Elías-Cabot, E., Raya-Povedano, J. L., Gubern-Mérida, A., Rodríguez-Ruiz, A., & Álvarez-Benito, M. (2022). Stand-alone use of artificial intelligence for digital mammography and digital breast tomosynthesis screening: A retrospective evaluation. *Radiology*, 302(3), 535–542.
42. Salim, M., Wählin, E., Dembrower, K., et al. (2020). External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. *JAMA Oncology*, 6(10), 1581–1588.
43. Samulski, M., Hupse, R., Boetes, C., Mus, R. D., den Heeten, G. J., & Karsssemeijer, N. (2010). Using computer-aided detection in mammography as a decision support. *European Radiology*, 20(10), 2323–2330.
44. Schaffter, T., Buist, D. S. M., Lee, C. I., et al. (2020). Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Network Open*, 3(3), e200265.
45. Sechopoulos, I., Teuwen, J., & Mann, R. (2021). Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Seminars in Cancer Biology*, 72, 214–225.
46. Siegel, R. L., Giaquinto, A. N., & Jemal, A. (2024). Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians*, 74(1), 12–49.
47. Spiesberger, W. (1979). Mammogram inspection by computer. *IEEE Transactions on Biomedical Engineering*, 26(4), 213–219.
48. Tabar, L., Fagerberg, C., Gad, A., et al. (1985). Reduction in mortality from breast cancer after mass screening with mammography. *The Lancet*, 1(8433), 829–832.
49. Tagliafico, A. S., Calabrese, M., Bignotti, B., et al. (2017). Accuracy and reading time for six strategies using digital breast tomosynthesis in women

- with mammographically negative dense breasts. *European Radiology*, 27(12), 5179–5184.
50. Tanyel, T., Denizoglu, N., Seker, M. E., Alis, D., Cerekci, E., Karaarslan, E., Aribal, M. E., & Oksuz, I. (2024). Mammographic breast positioning assessment via deep learning. arXiv:2407.10796.
 51. Van Winkel, S. L., Rodríguez-Ruiz, A., Appelman, L., et al. (2021). Impact of artificial intelligence support on accuracy and reading time in breast tomosynthesis image interpretation: A multi-reader multi-case study. *European Radiology*, 31(12), 8682–8691.
 52. Yala, A., Lehman, C., Schuster, T., Portnoi, T., & Barzilay, R. (2019). A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*, 292(1), 60–66.
 53. Yilmaz, E., Seker, M. E., Guldogan, N., Turk, E. B., Erdemli, S., Koyluoglu, Y. O., Sancak, S. N., & Aribal, M. E. (2025). Clinical application of AI in mammography: Insights from a prospective study. *Academic Radiology*.
 54. Yoon, J. H., Strand, F., Baltzer, P. A. T., et al. (2023). Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis: A systematic review and meta-analysis. *Radiology*, 307(5), e222639.