

1. AIを用いた乳腺画像診断の動向

井上 謙一 南大和クリニック乳がんセンター

近年、深層学習を中心とした人工知能 (AI) 技術の急速な進歩により、海外を中心に医用画像診断の分野において、AIや機械学習 (ML) を使用した AI/ML 搭載医療機器の開発・実用化が進んでいる。本稿では、日本や海外での現在の AI/ML 搭載医療機器の承認状況とその背景を中心に、今後の日本のあり方を考察していく。

承認モデルと米国の現実

現在、AI/ML 搭載医療機器の活用が最も進んでいるのが米国である。Singh らの研究¹⁾によると、1995～2024 年にかけて、Food and Drug Administration (FDA) が承認した AI/ML 搭載医療機器は 1016 件 (ユニーク機器としては 736 件) に上る。そのうち 84.4% が画像関連である。機能分類では、定量化・特徴局在化が 58% で最多だが、近年はトリアージや画像強調の比率が増加傾向にある。

FDA の 1000 件を超える承認の背景には、510 (k) 経路の活用がある。承認された AI/ML 搭載医療機器の大部分はこの 510 (k) という経路を通じて承認されているが、これは既存の機器と実質的に同等であることを示すことで承認を得るプロセスであり、必ずしも厳格な臨床試験を必要としない。Windecker らの承認された AI/ML 搭載医療機器 903 件を対象とした研究²⁾では、承認時点では 55.9% のみ臨床データを公開しており、28.7% のみ性別データを提示し、8.1% のみ前向き試験を実施し、2.4% のみ無

作為化試験を実施していた。多くは人口統計との比較や社会的背景も報告されていない。Muralidharan らの報告³⁾によると、3.6% のみが人種や民族を報告し、18.4% のみが対象者の年齢を報告していた。多様な集団での検証データが不足していると、少数派や特定の患者集団に対するバイアスや健康格差につながるリスクがある。

言い換えると、米国モデルは「ある程度の同等性を示せば市場に出し、実臨床での競争と検証によって性能を高める」というスタンスである。普及の速さは明らかだが、承認がそのまま臨床的に有効であることを保証するものでなく、汎化性能や人口動態の偏りが未検証のまま患者に届くリスクがある。

一方、日本の医薬品医療機器総合機構 (PMDA) は、臨床的有効性と安全性を確認した上で承認するスタンスを取る。例えば、筆者も参加していた DLADS 試験⁴⁾においては、マンモグラフィ画像から乳がんを検出する AI-CAD (computer-aided diagnosis) を開発した。後ろ向き試験で感度、特異度共に日本乳がん検診精度管理中央機構 (精中機構) の検診マンモグラフィ読影医師認定 B 判定に匹敵する精度を AI-CAD 単独で達成しているが、初のマンモグラフィ AI-CAD ということもあり、現在はプログラム医療機器 (Software as a Medical Device: SaMD) として承認を得るべく、前向き試験である MAJIC 試験を実施している。

米国と日本の制度の対比は単純な

優劣ではなく、どちらのリスクを社会が引き受けるかという制度設計の問題である。

日本で AI 乳腺診断が普及しない構造的要因

次に、AI/ML 搭載医療機器の構築自体が日本でなかなか進まない原因について考察していく。

1. データ問題：「データの島」

AI 開発の根幹はデータである。しかし、日本の医療機関は施設ごとに独立した「データの島」として存在しており、施設横断的なデータ収集は構造的に困難な状況にある。

第 1 の問題として、各医療機関からデータを収集すること自体が容易ではない。電子カルテシステムや画像管理システム (PACS) はベンダーごとに閉じた設計になっており、データの外部出力には多大な技術的・人的コストを要する。特に画像データは、1 件あたりのファイルサイズが大きく、加えて施設ごとのセキュリティポリシーの違い、システム担当者の人手不足、ベンダーとの契約上の制限なども障壁となり、研究目的であってもデータを収集する際は非常にコストがかかる。

2022 年に、内閣府の「経済財政運営と改革の基本方針 2022」などをベースとして医療 DX 推進本部が設置され、政府主導で「全国医療情報プラットフォーム」⁵⁾ 構想が推進されている。マイナン