



独立行政法人 医薬品医療機器総合機構
Pharmaceuticals and Medical Devices Agency

機械学習を活用する医療機器の審査について

AIを活用したプログラム医療機器専門部会の目的

AIを活用したプログラム医療機器開発を俯瞰し、課題を抽出することにより

- 現状の学問的なレビューではなく、①から③を中心に臨的に役立つ留意事項をまとめる
- 開発者への情報提供・PMDAにおける審査

に資することを目的とする。

論点

- ① データ再利用のあり方
- ② 評価データに求められる条件
- ③ 市販後に性能変化することを意図するAI (Adaptive AI) の審査のあり方

審査の考え方 - 承認拒否事由

- 医療機器の種類に応じた製造販売業の許可を受けていないとき
- 当該医療機器を製造する製造所が製造業の登録を受けていないとき
- 申請品目の審査の結果、次のいずれかに該当するとき
 - ① 効果又は性能を有すると認められないとき
 - ② 効果又は性能に比して著しく有害な作用を有することにより、使用価値がないと認められるとき
 - ③ 医療機器として不相当と厚生労働大臣が定めたものに該当するとき。
- 医療機器の製造所における製造管理・品質管理の方法が、QMS基準に適合していると認められないとき

(薬機法第23条の2の5)

承認審査のポイント：有効性と安全性のリスクベネフィットバランス

臨床^上でもたらされる
有効性

ベネフィット



リスク

臨床現場に導入された際に
発生し得るリスク

根拠資料とともに説明！

医用画像診断支援ソフトウェア（CAD）の評価の一例

開発コンセプト

放射線科医の見落としを減らす

臨床的有用性

臨床的なアウトカムに
貢献するか

例：
CADをつかった読影
と
CADを使わない読影
の診断成績を比較し
成績が向上すること

設計コンセプト

放射線診断専門医相当の検出能力

臨床性能

臨床データに対して
適切な性能を発揮できるか

例：
CADに臨床データを
入力した際の感度・特異度

基本的な性能

意図した仕様通りに
製品実現できているか

例：
実装した機能の
基本的な動作検証
ベンチテスト



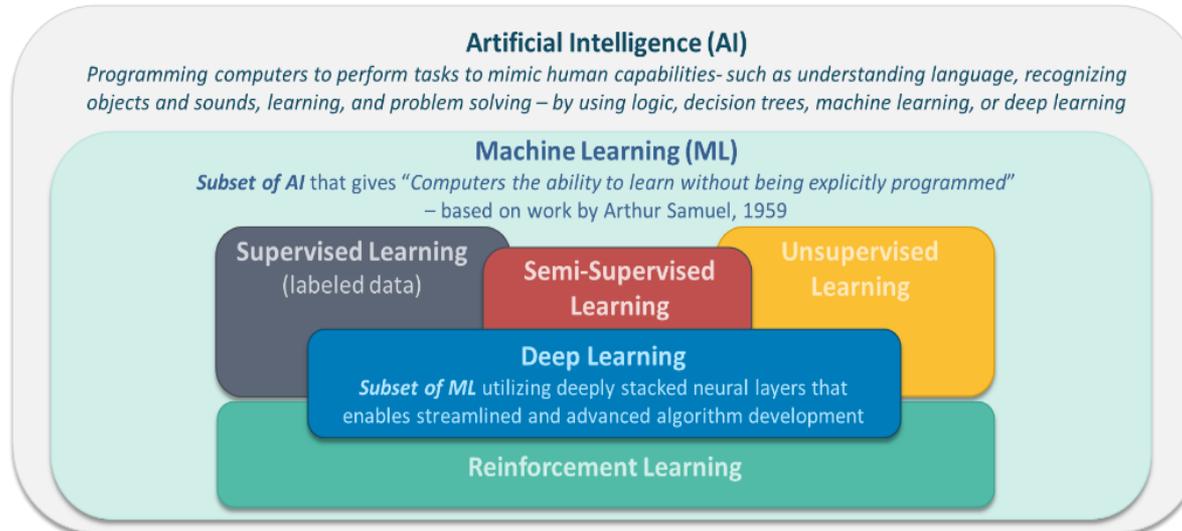
本邦で承認された機械学習を活用する医療機器

令和4年3月末現在

No.	承認日（一変承認日）	販売名	製造販売承認を受けた者
1	H30.12.6	内視鏡画像診断支援ソフトウェアEndoBRAIN	サイバネットシステム株式会社
2	R1.9.17	医用画像解析ソフトウェアEIRL aneurysm	エルピクセル株式会社
3	R1.12.25	類似画像症例検索ソフトウェアFS-CM687型	富士フイルム株式会社
4	R2.4.27	内視鏡画像診断支援ソフトウェアEndoBRAIN-UC	サイバネットシステム株式会社
5	R2.5.8	肺結節検出プログラム FS-AI688型	富士フイルム株式会社
6	R2.6.3 (R2.8.11)	COVID-19肺炎画像解析AIプログラムInferRead CT Pneumonia ※一変時に販売名を変更	株式会社CESデカルト
7	R2.6.19	AI-Radコンパニオン	シーメンスヘルスケア株式会社
8	R2.6.29 (R3.3.29)	内視鏡画像診断支援プログラムEndoBRAIN-EYE	サイバネットシステム株式会社
9	R2.6.29	COVID-19肺炎画像解析プログラムAli-M3	株式会社MICメディカル
1 1	R2.8.20	医用画像解析ソフトウェアEIRL X-Ray Lung nodule	エルピクセル株式会社
1 2	R2.9.2	内視鏡検査支援プログラムEW10-EC02	富士フイルム株式会社
1 3	R2.11.24	乳がん診断支援プログラムRN-デカルト	株式会社CESデカルト
1 4	R2.11.30	WISE VISION 内視鏡画像解析AI	日本電気株式会社
1 5	R3.5.26	COVID-19肺炎画像解析プログラム FS-AI693型	富士フイルム株式会社
1 6	R3.7.7	胸部X線画像病変検出(CAD)プログラム LU-AI689型	富士フイルム株式会社
1 7	R.3.9.1	肋骨骨折検出プログラム FS-AI691型	富士フイルム株式会社
1 8	R3.10.11	画像診断支援ソフトウェアKDSS-CXR-AI-101	コニカミノルタ株式会社
1 9	R3.12.9	胸部X線肺炎検出エンジン DoctorNet JLK-CRP	株式会社ドクターネット
2 0	R3.12.24	HOPE LifeMark-CAD 肺炎画像解析支援プログラム for COVID-19	富士通Japan株式会社

機械学習を活用する医療機器の定義

- 薬機法上、人工知能及び機械学習に関する定義はない。
- 国際医療機器規制当局フォーラム（IMDRF）における定義※は以下のとおり。
機械学習を活用する医療機器（MLMD ; Machine Learning-enabled Medical Devices）
意図する医療目的を達成するために、部分的又は全体として、機械学習を使用する医療機器



※ The descriptions within the diagram are not definitions, and are included to convey a general sense of the technology.

※ IMDRF/AIMD WG (PD1)/N67:2021

機械学習を活用する医療機器における臨床データの種別

臨床データに対する性能試験

臨床データに対して適切な性能を発揮できるか確認するための試験（学習アルゴリズム及び学習モデルの評価）

臨床データ

- A) 学習データ※：機械学習モデルを学習するために使用されるデータ
- B) バリデーションデータ※：ハイパーパラメータの調整又は選択するアルゴリズムの検証に使用されるデータ
- C) 評価データ※：最終モデルの性能を評価するために使用されるデータ



臨床上の有用性を評価する性能試験

臨床的位置づけを加味した臨床的アウトカムに貢献するか確認するための試験。

- C) 評価データ※；最終モデルの性能を評価するために使用されるデータ

- 審査ではどのようなデータで開発されたか確認するが、妥当性を問うのは評価データ。
- どのような評価データに対しどのような結果が得られたかが性能となり、MLMDはブラックボックスの特性からこの結果が重視される。
- 実臨床を反映した評価データであることが望ましいが、どの程度バリエーションが必要かは難しい。

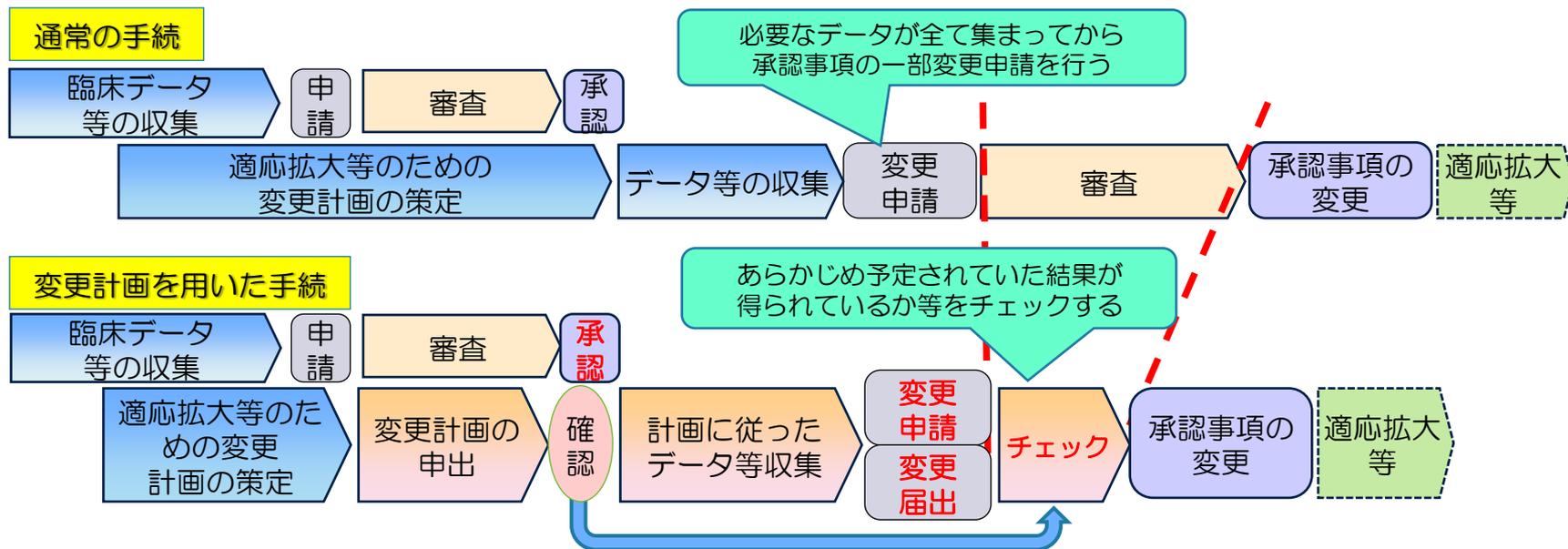
可塑性の有無からみた機械学習を活用する医療機器の種別

- Locked type : 機能が変更しないようにアルゴリズムを固定
 - 通常の審査と基本コンセプトは同様。
- Adaptive type : 新たなデータに基づき、繰り返しその機能等を変更
 - 初回の性能の評価は通常の審査と同様。
 - 医療機器の特性に応じ将来改良が見込まれている医療機器について、その改良計画自体を承認する制度（医療機器の特性に応じた変更計画の事前確認制度：IDATEN制度）が令和2年9月に施行された。
 - 一般に、IDATEN制度の利用が期待される。
- なお、Locked type 及び Adaptive type は、IMDRF、ISO/IEC等を含め明確な定義は存在しない。

医療機器の変更計画の確認制度 (IDATEN)

■法第23条の2の10の2第6項■

第1項で確認を受けた「変更計画」に従った変更を行う日の省令で定める日数前までに届出たときは、承認を要しない。



IDATEN制度の運用にあたり以下事項が想定される

- 市販後の製品の改良等において、初回の承認申請に用いた評価データの再利用
- 市販前に使用していた機械学習アルゴリズムの変更、同一の機械学習モデルにおけるハイパーパラメータの変更

論点① データ再利用のあり方

- 市販後の製品の改良等において、初回の承認申請で用いた評価データを再利用することが想定される。
- 評価データの収集やアノテーション作業は負荷が大きく、独立した評価データを都度用意することはコストが高いことから、再利用のニーズは高い。
- 一方で、評価データの再利用する場合、改良プロセスと評価データを適切に分離しないと、（意図せず）単に評価データに対する最適化がなされてしまい、汎化性能を損なっているにもかかわらず気づけない可能性がある。

<対応例>

- 評価データの再利用を前提とした、モデルの改良プロセスの規定。
- 評価データとは独立した、小数の別のデータから汎化性能を確認する。
- 評価データの利用回数に上限を規定する。

データ再利用の可否や再利用するための要件とは何か

具体事例

販売名：n o d o c a（ノドカ）、製造販売業者：アイリス株式会社 承認番号：30400BZX00101000
審査報告書一部抜粋

https://www.pmda.go.jp/medical_devices/2022/M20220516002/112714000_30400BZX00101_A100_3.pdf

2) 治験機器の改良に関する妥当性

総合機構は、本品の開発の中で行われた治験機器の改良内容に関する妥当性について、本臨床試験結果から意図的な性能向上が本品で行われていないことを申請者に説明を求めた。

申請者は以下のように説明した。

本臨床試験成績を [REDACTED] 製品
[REDACTED] 状
態とした。また、本臨床試験データの管理にあたっては、 [REDACTED]
[REDACTED] を担保できるような開発プロセスと
した。

総合機構は、申請者の説明から、本臨床試験結果に基づく意図的な製品改良はできない状況であったと考え、当該説明は受入れ可能と判断した。

具体事例

販売名：n o d o c a（ノドカ）、製造販売業者：アイリス株式会社 承認番号：30400BZX00101000
審査報告書一部抜粋

https://www.pmda.go.jp/medical_devices/2022/M20220516002/112714000_30400BZX00101_A100_3.pdf

3) 追加試験成績から本品の有効性を評価することの妥当性について

総合機構は、追加試験成績を主たる評価資料とすることの妥当性について、本品の改良が本臨床試験結果のみに最適化されたものではないことを申請者に説明を求めた。

申請者は以下のように説明した。

本臨床試験の患者データ以外のデータとして、臨床研究（jRCTs03219120）にて収集された患者データを用いた試験成績（以下「臨床研究データを用いた試験成績」という。）を提示する（表 21）。この臨床研究データを用いた試験は、臨床研究（jRCTs03219120）にて収集された患者データのうち、XXXXXXXXXX PCR 法検査陽性 XXXX 例、陰性 XXXX 例の患者データを用いて実施した。その結果、感度 XXXX%、特異度 XXXX%であり、追加試験の成績（感度 76.0%、特異度 88.1%）と臨床研究データを用いた試験成績（感度 XXXX%、特異度 XXXX%）を比較し、両者において同等の性能であることが認められたことから、本品の改良が本臨床試験結果のみに最適化されたものでないとする。

総合機構は、本臨床試験及び臨床研究データを用いた試験成績の両者において、同等の性能であることが認められたことから、本臨床試験結果のみに最適化された改良ではないと考え、専門協議の議論も踏まえ、追加試験にて本品の有効性評価を行うことは妥当であると判断した。

論点② 評価データに求められる条件

- 機械学習を活用する医療機器は、ブラックボックスの特性より、開発者が認知し得ない弱点が存在し得る。
- したがって、その製品が臨床上想定される様々なバリエーションのデータを網羅的に収集し評価することで、明らかな弱点がないかを確認する必要がある。
- 一方で、網羅的であることを客観的に説明することは困難であり、また際限なくデータを収集することは現実的ではない。

<評価データのバリエーション例>

- ベンダー及び撮像条件
- 適用患者
- 疾患・病変位置
- 画像取得施設
- 実際の有病率の陽性・陰性画像セット

**評価用データのバリエーション
の考え方や十分性
また評価データが存在しない
パターンに対する対策の考え方
について**

論点③ 市販後に性能変化することを意図するAI (Adaptive AI) の審査のあり方

- 市販後の学習として医療施設毎によるバージョンアップが想定されるが、製造販売業者と同様の品質マネジメントを臨床現場に求める工夫等が必要と考えるが、より具体的な事例について検討いただきたい。

<対応例>

- 臨床上許容できる性能を下回るときに初期値に戻す機能
- 臨床的位置づけが変わらないように事前に性能を固定する機能

- 市販後の学習を通じて、市販前に使用していた機械学習アルゴリズムの変更、同一の機械学習モデルにおけるハイパーパラメータの変更が想定されるが、より具体的な事例について検討いただきたい。

<具体事例>

- サポートベクタマシンから深層学習への変更
- 深層学習における層数の変更

一品目の範囲に関する整理は、当該論点の対象外。