

人工知能等の先端技術を利用した医療機器プログラムの薬事規制のあり方に関する研究

中岡 竜介¹, 清水 昭伸², 野村 行弘^{3, 4}, 花岡 昇平³
西川 和宏⁵, 古川 浩⁶, 薮島 由二¹

¹国立医薬品食品衛生研究所 医療機器部

²東京農工大学大学院 工学研究院

³東京大学医学部附属病院

⁴千葉大学フロンティア医工学研究センター

⁵日本メジフィジックス株式会社

⁶(一社)日本画像医療システム工業会

人工知能(AI)を利用した医療機器プログラムの急速な発展
↓
市販後性能変化も含め、従来の薬事では対応できないケースが生じる

開発を妨げず、安全性・有効性を迅速に評価するための薬事規制システムを検討する必要性

次世代事業等で抽出されている課題例

従前のアルゴリズムの延長にもかかわらず、「AI」に対する過度な期待
→少なくとも産官学での共通理解が必要



人間が行える作業は全てAIに取って代わられるかも？



まだ、教えられたことしかできないからムリムリ！！

人の代わりにはならない(限界がある)

深層学習で構築したアルゴリズム評価の難しさ

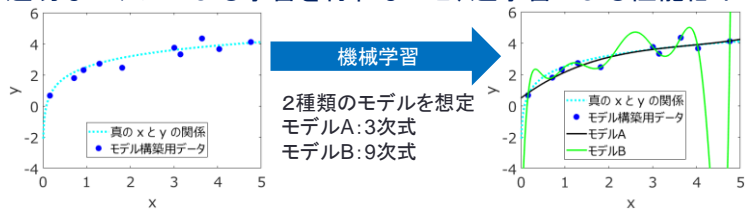
→適切な性能評価での代替が必要

- 多数のパラメータを持つ大量のデータを処理可能だが、訓練によりモデルを構築するため、最終的なモデルが分からない
- 学習データに含まれていない結論は予測困難
- 100%の正しさは保証不可能(本質は確率論) 等

学習データ、基本モデル、テストデータ等の妥当性評価の難しさ

→目的に応じた適切なデータ、モデル確保手段の構築が必須

例)適切なモデルによる学習を行わないと、過学習による性能低下が生じる



モデルBは全てのデータを通過するように適合しているが、それ以外では誤った結果を示す
引用元: <https://datachemeng.com/overfitting/>

本研究の目的:
AMED研究事業による解決策の検討と将来あるべき薬事規制システムに関する提言を作成

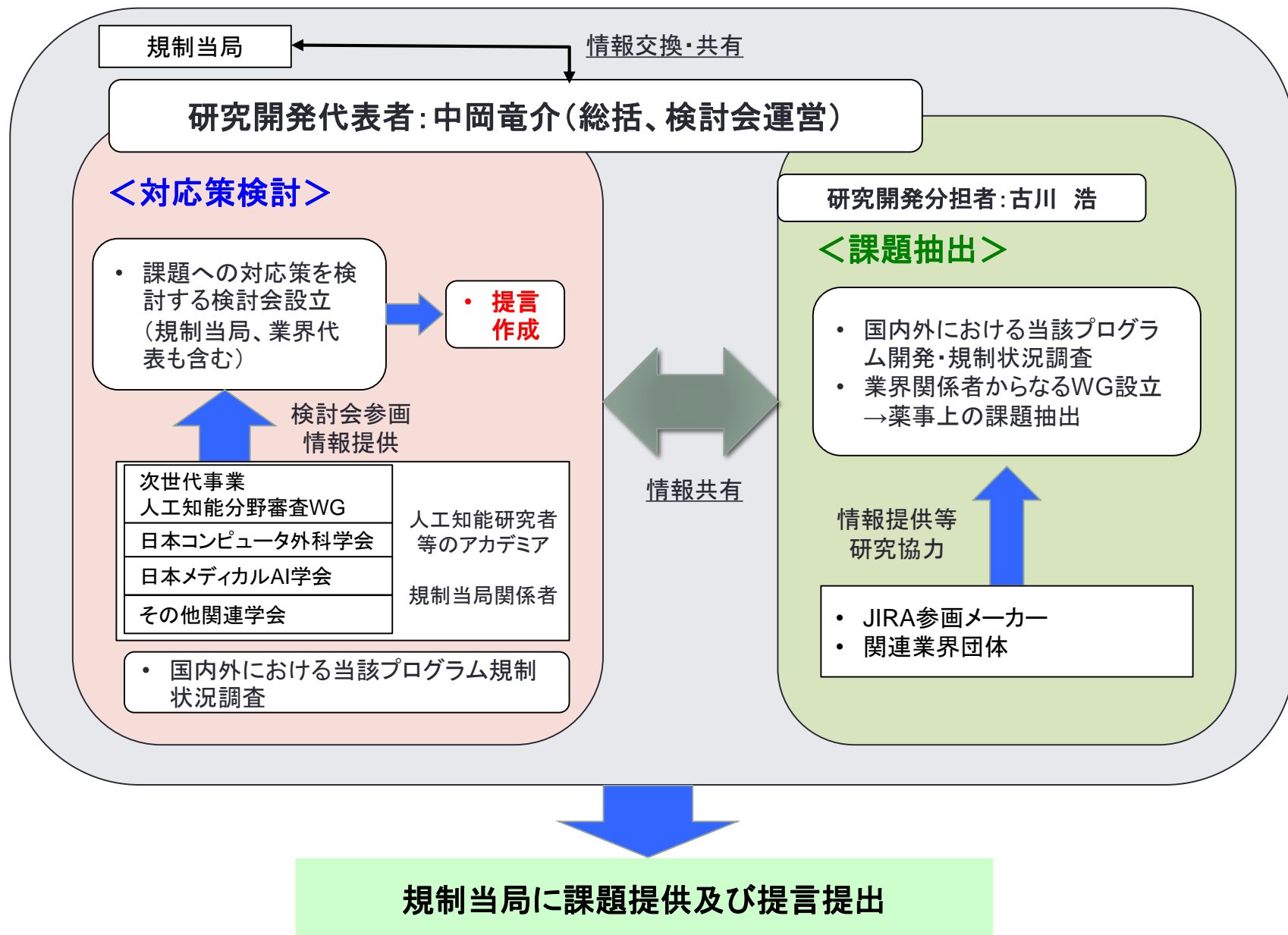
他にも想定される様々な薬事的課題を産官学(衛研、業界団体中心)で抽出・整理

抽出・整理された課題についての解決策等を有識者で議論

議論を経て取りまとめられた解決策等を整理し、提言として適宜規制当局に提出

アカデミア関係の有識者としては、人工知能分野審査WG委員経験者、日本メディカルAI学会及びコンピュータ外科学会等からの推薦者が参画

対象となる技術の革新速度は想定以上に著しいことから、課題抽出は3年間継続的に行い、それに対する解決策を適宜提言に取りまとめて規制当局に提出する



これまでにWGで抽出された課題



対象	想定される課題
AI 医療機器	<ul style="list-style-type: none"> 分類（診断、治療、予防、体の構造機能に影響を及ぼすなど）と定義 設計過程のみでAI機能を使用した機器の扱い 保守ツールとしてAI機能を使用した場合（広告との関係が中心）
研究・製品開発に用いるデータの法的な扱い（要配慮個人情報の扱い）	<ul style="list-style-type: none"> 市販後学習における個人情報の扱いの考え方 企業における研究行為に関する個人情報の扱い 医療機関、学会（医師）による研究行為に関する個人情報の扱い 利用可能なデータの明確化（収集済みのデータの扱い等を含む）
学習データ	<ul style="list-style-type: none"> 必要なデータ数等に関する考え方の確認 製品化に使用したデータの保存期間について 外国データの使用について（米国、欧州等） データの特定化（装置、撮影条件等）及び動画の扱いについて
検証データ	<ul style="list-style-type: none"> 検証データの信頼性（信頼性調査の扱い）
教師付きデータ	<ul style="list-style-type: none"> 教師データ作業に関する資格要件は、最終確認を医師がすればよいか？ 医師に関する要件は？ 日本の医師資格が必須か？ 外国の医師資格の扱いについて検討
学習済モデルの扱い	<ul style="list-style-type: none"> 他社開発モデルの扱いは（設計管理の有） 他社販売のモデルの扱いは（設計管理の有無等） 医師の研究成果の扱い
医療機器の使用目的に伴うリスクに関する事項	<ul style="list-style-type: none"> 人に対する直接のリスクの扱い（現状の直接のリスクの考え方でよいか） 結果の使用が与える間接リスクを検討する必要があるか？（CADxの扱いにはこの概念が入っていると思われる） First reading concurrent に関する扱いの整理
承認申請に関する事項	<ul style="list-style-type: none"> 承認、認証区分の整理 市販後学習により性能が変化する承認書記載事項 学習に伴う性能変化と一変等手続き 記載事例（申請パッケージ）

青字：アカデミアによる検討会での討議を経て成果物を整備済み
 黄色マーカ：調整費事業による検討対象（検討結果を基に討議）

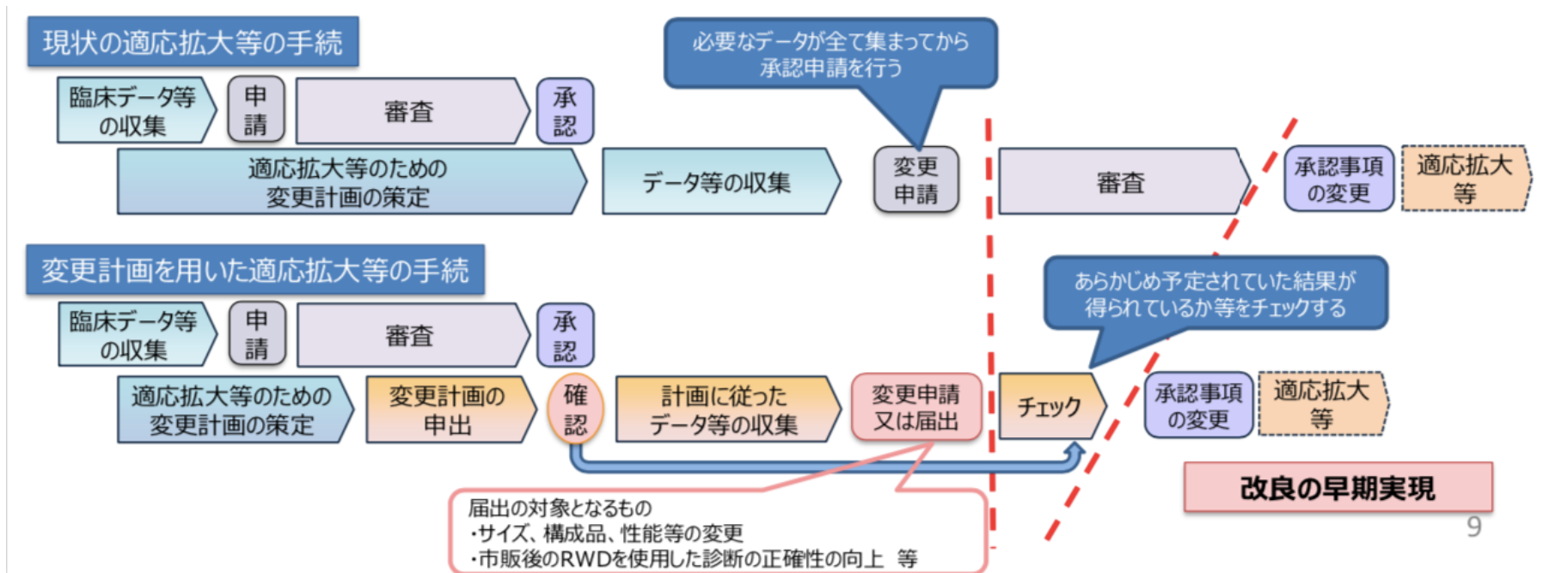
調整費による追加研究事業について

令和元年度薬機法改正(2019.11成立)



医療機器の継続した改良を可能とする承認審査制度が導入(2020年秋より施行)
Improvement Design within Approval for Timely Evaluation and Notice (IDATEN)

改良が見込まれている医療機器について、変更計画を審査の過程で確認し、計画された範囲の中で迅速な承認事項の一部変更を認める制度



AIを活用した医療機器のように市販後に性能等が変化し得る医療機器も対象だが、予定通りに性能を向上させることは可能か？

妥当な性能向上計画を設計・実現させるため、規制・開発の双方にとって性能変化に影響を及ぼす要件やその妥当性検証方法を明らかにし、共通認識とする必要性が生じている

IDATENのAI利用医療機器への適用実現へ



市販後の追加学習(市販後学習)に伴うAIの性能変化に対する具体的な薬事戦略を先駆けて構築することを目的とした活動が始まっている(市販後学習を意図した製品に対する対策)
→市販後学習による性能変化に関する科学的知見の蓄積が根拠として必要になる

米国の状況

- 2019年4月、市販後学習を可能とした医療機器の実現化に向け意見募集を目的としたDiscussion Paperを発行
- 2020年2月、FDA主催のPublic Workshopが開催され、Discussion Paperに対するリアクションとして産学が意見を発表、またFDAが2021年に現時点での考え方を発表しており、産官学での意見集約と共通認識構築を進めようとしていることが明らかになっている



目指す規制方向は日本と同様だが、参考情報としては個別研究を取り上げた比較検討が示されているのみであった→データの質が性能に及ぼす影響を比較した系統的検討がない

米国に遅れを取らないために、AIの性能変化に影響を及ぼす要因を明確にするための系統的な実証研究を行い、その結果を速やかに規制(IDATEN)に反映させることが望まれる

段階的とはいえ、製造販売業者による市販後学習で性能変化が可能なAI利用医療機器の社会実装をいち早く実現することが期待され、規制・開発の双方に貢献できる

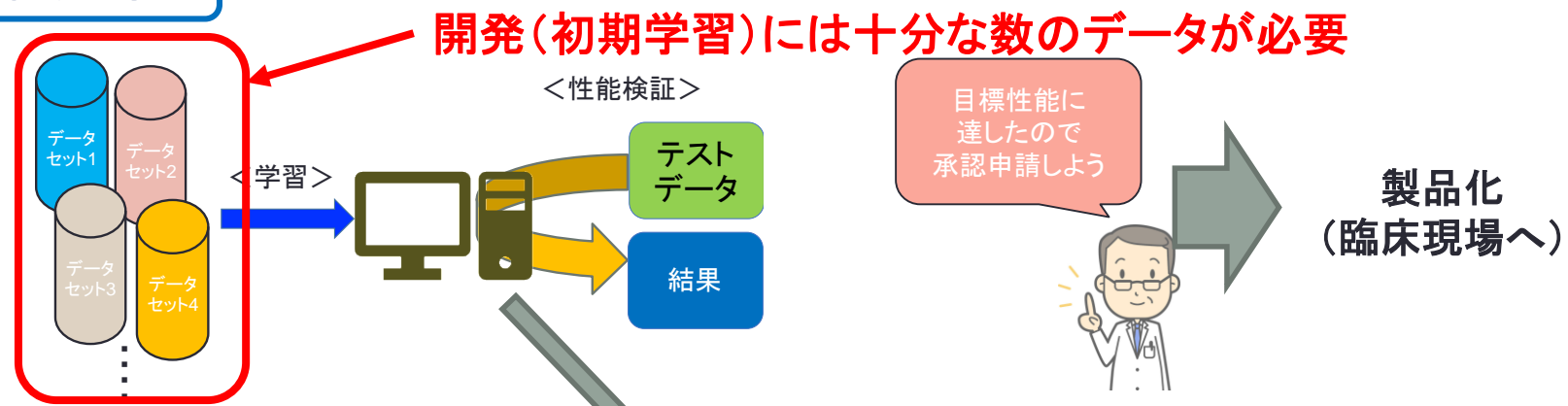
- 機械学習を利用した場合、特徴の一つである継続的な学習に伴う性能変化を実装するために解決すべき課題が多い
 - 使用施設毎に性能が変化してしまうため、「医療機器」としての承認時における品質、性能を担保することが可能か？
(承認時と比較して性能は変化し得る製品を一つの製品と扱えるか？)
 - 学習に使用するデータに対して要求すべき品質が不明
 - 製品開発時のデータとどの程度の同等性が必要になるのか？
 - 異なる装置、パラメータで撮像したデータは市販後学習に使用できるのか？
 - 学習後に変化した性能に対する評価を承認時と同じ方法で実施するのは妥当か？
 - 適応拡大でなくても、テストデータが対象とする疾患の母集団を代表しているか分からないし、その妥当性をどのようにすれば示せるのか？

性能変化に影響する市販後学習データの要因、変化した性能を適切に評価できる性能評価プロセスの要件等の明確化とガイダンス等の作成により、(IDATENを含めて)承認手続き等を円滑に運用する体制を構築、製造販売業者による段階的な市販後学習を可能とする
(製造販売業者による市販後学習は既にIDATENの適用が可能)

知見を蓄積し、機械学習の特徴を活かした「施設毎の市販後学習」や「継続的な市販後学習」機能を有するAI利用医療機器プログラムの実用化へと繋げる

AMED調整費による検討ポイント

AI開発時



調整費事業

市販後学習を行ったAI利用医療機器の性能に影響を与える要因の明確化と薬事規制への提言が目的 (中立的な機関が中心となって行うべき研究(RS研究での実施が望ましい))

既存のAI利用医療機器プログラムを活用し、市販後学習を実施

(開発における初期学習と異なり、必要となるデータは検証対象要因等を考慮して収集する必要がある)

利用予定のプログラム(参画機関が製品化に向けて開発中、あるいは製品化したプログラム)

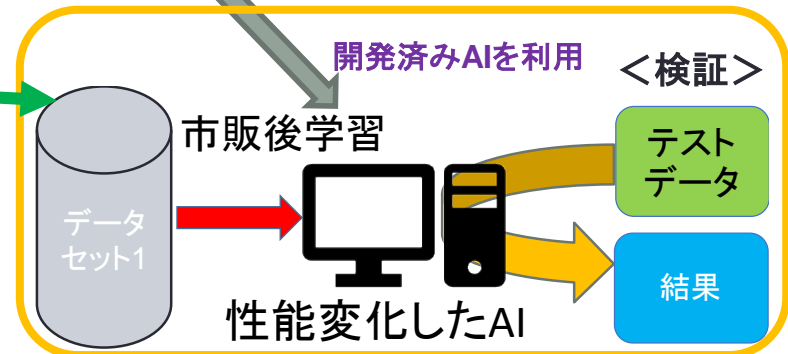
- ・東京大学医学部附属病院: MRA画像から脳動脈瘤を検出するプログラム
- ・東京農工大学、日本メジフィジクス: 骨シンチ画像からがんの骨転移指標を提示するプログラム

市販後学習に必要な追加学習データ

- ・MRA画像
- ・骨シンチ画像

- 検討する検証対象要因に特化し、且つ初期学習時とは異なるデータが必要
- アノテーションを作成する医療従事者の特性の違いの影響を検討するためには、既存の画像に対する追加作業が必要
- いずれのデータも、現在、データベース化されていない etc.

→新規収集が必要



新たな研究体制（実証研究グループの追加）



薬事的提言に関する研究グループ

研究開発分担者
所属：（一社）日本画像医療システム工業会（JIRA）
氏名：古川 浩

- AI等を利用した医療機器プログラムの薬事規制における課題抽出に関する研究

情報共有

研究開発代表者
所属：国立衛研
氏名：中岡竜介

- 総括及び産官学連携体制構築
- AI等を利用した医療機器プログラムの薬事規制における課題への対応策検討に関する研究（提言作成を含む）
- 受診患者群の違いに着目したAIを用いた骨シンチ画像解析医療機器プログラムの市販後学習後に生じる性能変化に関する研究

厚生労働省
PMDA

情報交換
提言提出

検討結果等提供

JIRA参画企業
関連業界団体
AI分野専門家
人工知能分野審査WG参画委員
コンピュータ外科学会
日本メディカルAI学会

WG、検討会等参画、研究協力

実証研究グループ（調整費）

研究開発分担者
所属：東京農工大学
氏名：清水 昭伸

- 画像の質のばらつきに着目したAIを用いた骨シンチ画像解析医療機器プログラムの市販後学習後に生じる性能変化に関する研究

研究開発分担者
所属：東京大学医学部
附属病院

氏名：野村 行弘

- AI技術を用いたCT/MR画像用医療機器プログラムの市販後学習における性能変化に関する研究：正解入力作業の影響について

研究開発分担者
所属：日本メジフィジックス株式会社
氏名：西川 和宏

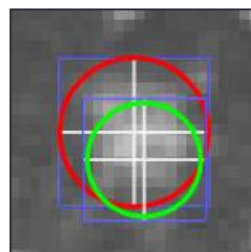
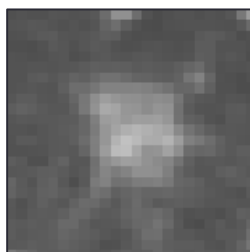
- AIを用いた医療機器プログラムへの市販後学習に必要なデータのAnnotation作業に関する研究

アノテーションの影響

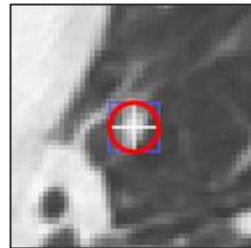
東大病院グループ担当

- 正解入力作業に携わる放射線科診断医の能力及び人数による市販後学習後の性能変化検証および影響要因候補の明確化

2名の放射線科医による正解入力のばらつきの例



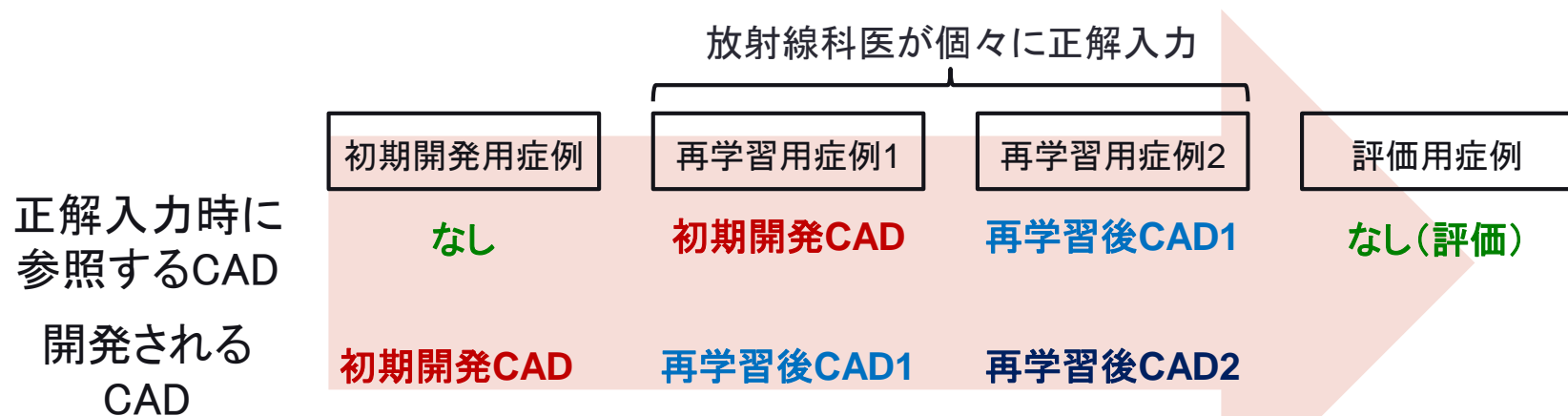
2名とも入力しているものの、入力範囲は異なる



1名のみ入力

赤：放射線科医1、緑：放射線科医2

- 対象CADソフトウェア（東大病院でモデル開発済）
 - 胸部CT画像の肺結節検出
 - 頭部MRA画像の脳動脈瘤検出
- 正解入力・再学習の手順



初期開発用症例、評価用症例: 別途、複数名の放射線科診断専門医により正解を入力

- 評価指標

- FROC曲線の2FPs/caseにおけるsensitivity[%]

肺結節検出を目的としたモデルプログラムの場合

- Annotation実施者に応じて再学習後に変化した性能にはばらつきが認められた（初期開発時より性能が低下する場合も認められた）
- Annotation実施者の経験年数と性能変化方向との関連性は見出せなかった
- 脳動脈瘤検出を目的としたモデルプログラムを用いた場合でも、同様の結果となった



複数名によるAnnotationを模したデータを用いて同様の検討を実施

現在論文執筆中のため、詳細なデータは非公開

撮像装置及び条件の影響

国立衛研、東京農工大学

日本メジフィジックス(株)担当

- 診断支援システムの性能

- システム使用施設のデータを用いた追加学習で性能が向上する可能性がある
- データに内在する要因が追加学習後の性能に影響する可能性がある



具体的な要因を検出する必要がある

機械学習を用いて開発されたモデルを用いて市販後学習を模した追加学習を行い、得られた性能変化から、撮像装置及び条件（臨床施設に応じた条件）が与える影響を検討する

モデル医療機器プログラム:

骨シンチグラムにおける陽性高集積検出システムの基礎モデル
（当該システムは東京農工大学が開発し、臨床応用されている）

骨シンチデータ: 日本メジフィジックスが臨床機関から収集後、アノテーション処理を実施

複数の施設から所定の手続きを経て骨シンチ画像を入手



骨転位部分の正解ラベル付与(アノテーション)処理を日本メジフィジックスが担当



得られた画像を用いて初期学習及び市販後学習を模した追加学習を実施



収集画像の一部を予めAIにとって未知のテストデータとして確保しておき性能検証を実施

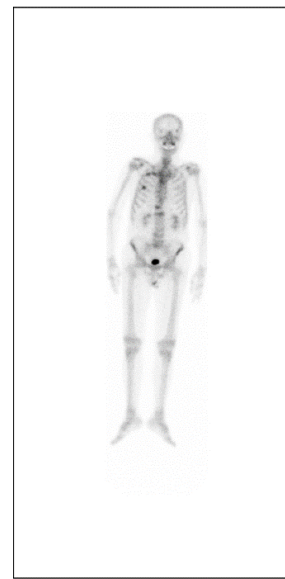
施設	症例数	画像枚数
000	246	492
001	160	320
002	156	312
003	173	346
004	158	316
005	139	278

実験用データセットの内訳

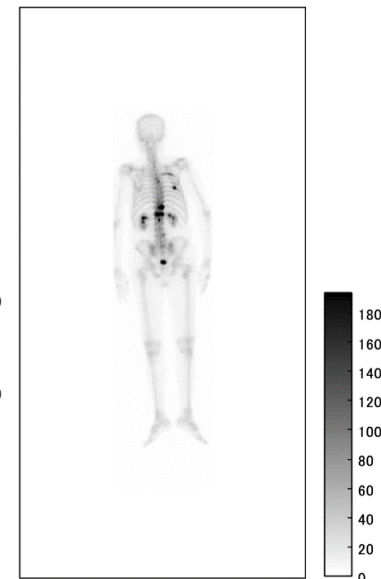
骨シンチグラムの例

前面

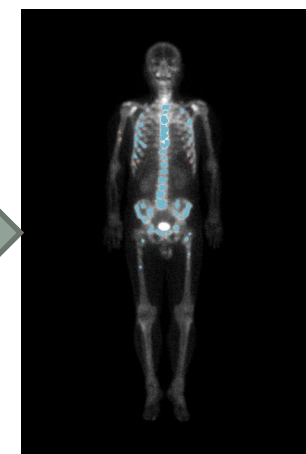
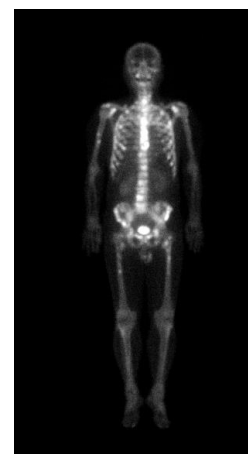
後面



(a)



(b)



- ・初期AIモデル構築において使用するWorkstationが異なると最終的な性能が異なる可能性があること、保存するモデル数を増やすことで性能差が小さくなることが明らかとなった(市販後の追加学習においても同様の現象が起こることが予想されるため、考慮すべき要因の一つになる)
- ・最適なモデルを構築し、市販後学習を模した追加学習を施設毎のデータで実施したところ、その性能変化方向は様々であり、データの各種パラメータが影響を及ぼしていることを示唆するデータを得た

現在論文執筆中のため、詳細なデータは非公開

- ・市販後学習後の性能に影響を及ぼす影響の明確化を目指し、以下の要因に着目した比較検討を東京農工大学との共同研究により継続して実施する
 - 撮影条件
 - 受診患者群
 - 学習症例数
 - 学習対象となるWorkstationの違い 等
- ・市販後学習時の医療機器プログラムにおける性能変化を適切に評価可能なプロセス構築に資する実証研究を実施する
 - 市販前テストデータを用いた市販後モデル性能評価の妥当性検討
 - テストデータ再利用に関する検討
 - Underspecification問題への対応

研究内容を基に、レギュラトリーサイエンスに立脚したガイダンスを作成し、先行研究で提出した提言と併せて当該医療機器の承認審査における適切な体制構築と当該プログラムの社会実装に貢献する

提言等作成

<検討会メンバー>

(座長) 橋爪 誠(九州大学／北九州中央病院)
(委員) 清水 昭伸(東京農工大学)
中野 壮陞(医療機器センター)
浜本 隆二(国立がん研究センター研究所)
堀尾 貴将(森・濱田松本法律事務所)
森 健策(名古屋大学)
山田 真善(国立がん研究センター中央病院)
横井 英人(香川大学)

【連携学会】

日本コンピュータ外科学会
日本メディカルAI学会

【厚生労働省】

医薬・生活衛生局 医療機器審査管理課

【PMDA】

医療機器審査第一部、第二部
プログラム医療機器審査室

<WGメンバー>

(委員) 板垣 直樹(MT JAPAN)
伊藤 毅 (JIRA)
小掠 真貴(NEC)
桜井 智也(分析工)
小掠 真貴(NEC)
館野 誠(日補工)
手塚 智(JEITA)
中光 敬(JAHIS)
羽田 健太郎(JEITA)
松川 智彦(日医工)
水野 米治(眼医器協)
村上 邦臣(歯科商工)
吉田 眞司(日医光)

(事務局: JIRA)

船橋 毅
関口 敬仁
田中 弘

その他、医療機器業界団体関係者

実証実験

東京農工大学: 島田夏帆、柴程軻、藤本 啓輔
東京大学医学部附属病院放射線科、コンピュータ画像診断学/予防医学講座の皆様

ご静聴、ありがとうございました