

患者さんに寄り添う医療データの活かし方 —医療AI研究開発との共創—

国立がん研究センター研究所医療AI研究開発分野・分野長
一般社団法人日本メディカルAI学会・代表理事

浜本 隆二

人工知能（AI）を活用した統合的ながん医療システム開発プロジェクトの概要 （2016年開始）

2018年からは内閣府主導の官民研究開発投資拡大プログラム（PRISM）「新薬創出を加速するAIの開発がCRESTプロジェクトにアドオンされ、また2023年6月より内閣府BRIDGE事業「医療デジタルツインの発展に資するデジタル医療データバンク構想」が開始された。

 **内閣府**
Cabinet Office, Government of Japan

 **PRISM** 官民研究開発投資拡大プログラム
Public/Private R&D Investment Strategic Expansion Program

 **BRIDGE**
研究開発とSociety 5.0との橋渡しプログラム
programs for Bridging the gap between R&D and the Ideal society (Society 5.0) and Generating Economic and social value

 国立研究開発法人
国立がん研究センター
National Cancer Center Japan



新規がん診断システム

**人工知能[AI]技術
(機械学習・深層学習)**
AI技術を用いて大規模な医療データに潜在的に含まれている特徴量を抽出する

個別化医療
実現支援システム

新規創薬設計システム

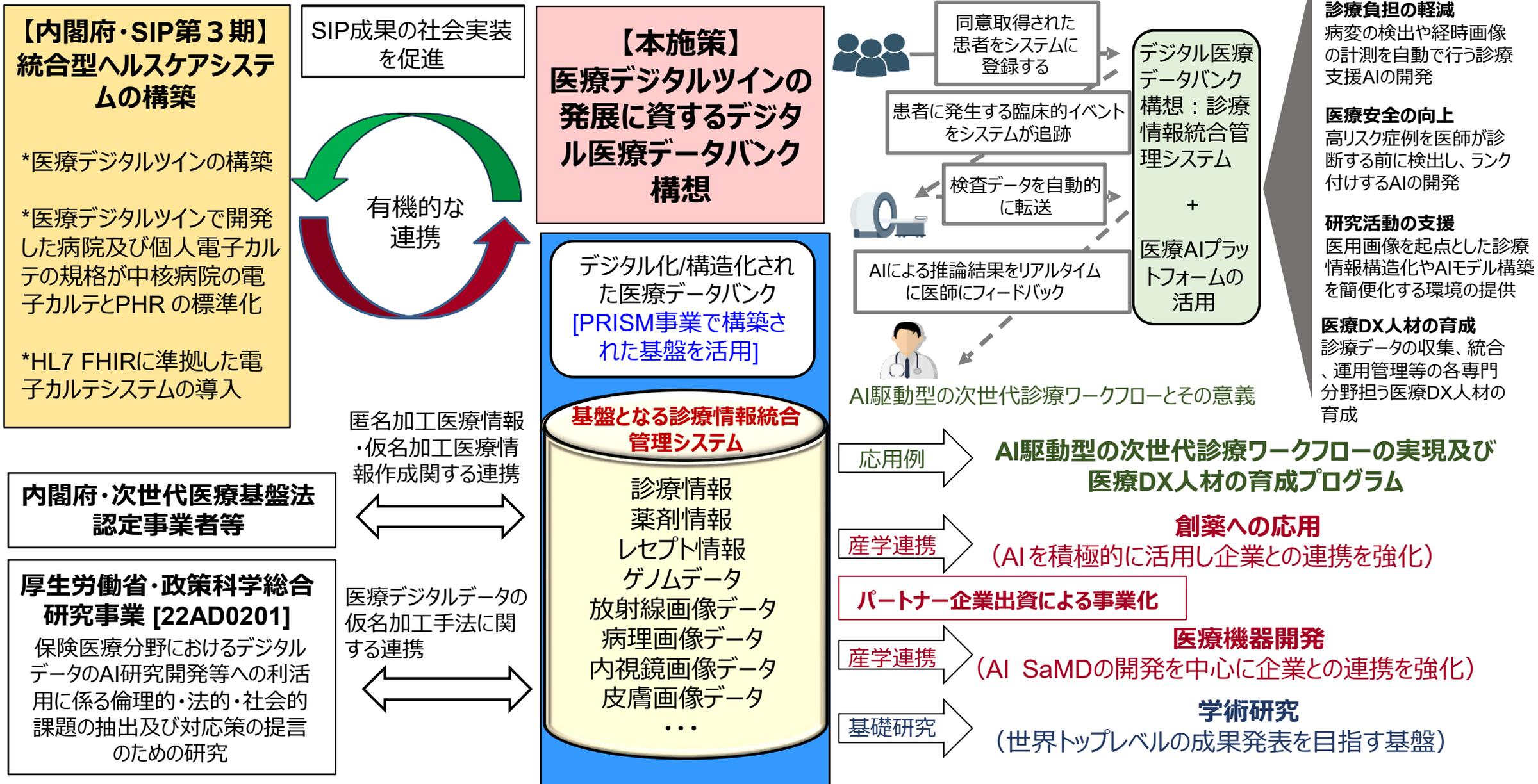
実臨床への応用を最重要視

医療費削減に貢献

ヘルスケア産業への展開

JST AIPネットワークラボ/CREST「AI」研究領域
研究代表：浜本 隆二

BRIDGE「医療デジタルツインの発展に資するデジタル医療データバンクの構築と社会実装」



「医療デジタルデータのAI研究開発等への利活用に係るガイドライン」の発出

事務連絡
令和6年9月30日

関係各施設等機関等
関係各国立研究開発法人
関係各独立行政法人
各都道府県 関係部局 御中
各特別区
各保健所設置市
関係各団体

厚生労働省大臣官房厚生科学課
厚生労働省医政局研究開発政策課

「医療デジタルデータのAI研究開発等への利活用に係るガイドライン」について

平素より厚生労働行政に御協力いただき厚く御礼申し上げます。

今般、令和5年度厚生労働科学研究費補助金・行政政策研究分野・政策科学総合研究（倫理的・法的・社会的課題研究事業）「保健医療分野におけるデジタルデータのAI研究開発等への利活用に係る倫理的・法的・社会的課題の抽出及び対応策の提言のための研究」（研究代表者：中野 壮陸（公益財団法人医療機器センター））の分担研究「デジタルデータのAI研究開発等への利活用に係るガイドライン作成班」（研究分担者：浜本 隆二（国立研究開発法人国立がん研究センター））において、別添のとおりガイドラインが作成されました。当該ガイドラインは、医療機関、学術研究機関及び民間企業等が共同研究を起点として、医療機関等が保有する医療情報を利活用した製品開発を行う場合を想定し、研究開発のステージに応じて、医療情報を利活用するための適切な法的根拠を明確化するとともに、医療情報の特性を踏まえた仮名加工情報の作成手順やその運用に関して取りまとめたものであり、医療デジタルデータと人工知能（AI）を研究開発等に利活用するに当たって留意すべき内容について記載しております。

つきましては、貴管下関係施設において、別添をご参照いただくとともに、貴管下関係者への周知方お願い申し上げます。

目次

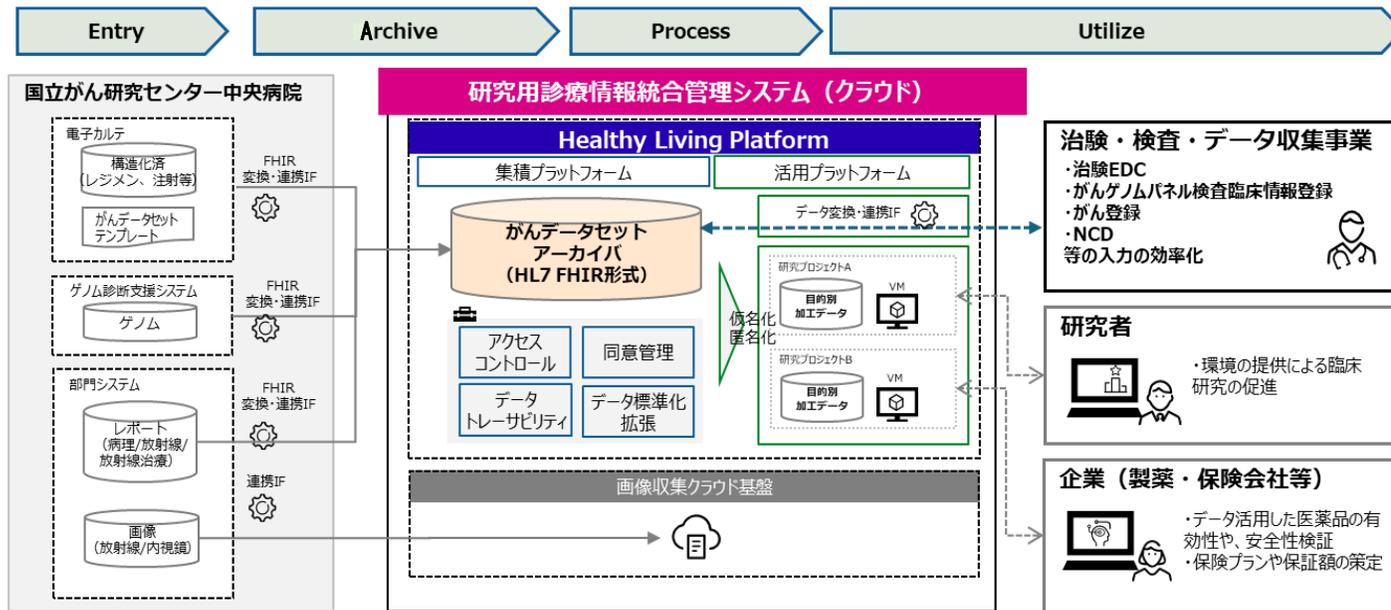
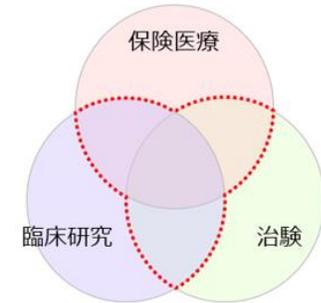
- はじめに
 - 背景
 - 目的と対象
 - 用語の定義
- 医療情報の利活用と法的根拠
 - 医療情報の利活用
 - 医療情報の利活用に関わる事業者
 - 要配慮個人情報としての医療情報に係る規律の概要
 - 学術研究例外とその該当性
 - 公衆衛生例外とその類型
 - 製品開発目的で医療情報を医療機関等から民間企業等に提供する際の法的根拠
- 仮名加工情報とその共同利用
 - 仮名加工情報とは
 - 仮名加工情報の共同利用
 - 仮名加工情報の共同利用の設定パターン
 - 推奨されない仮名加工情報の取扱い
 - 仮名加工情報の個人情報該当性
 - 仮名加工情報の規律
 - 仮名加工情報の利用目的変更・共同利用等に関する手続き
 - 削除情報等の安全管理措置義務

デジタル医療データバンクの構築

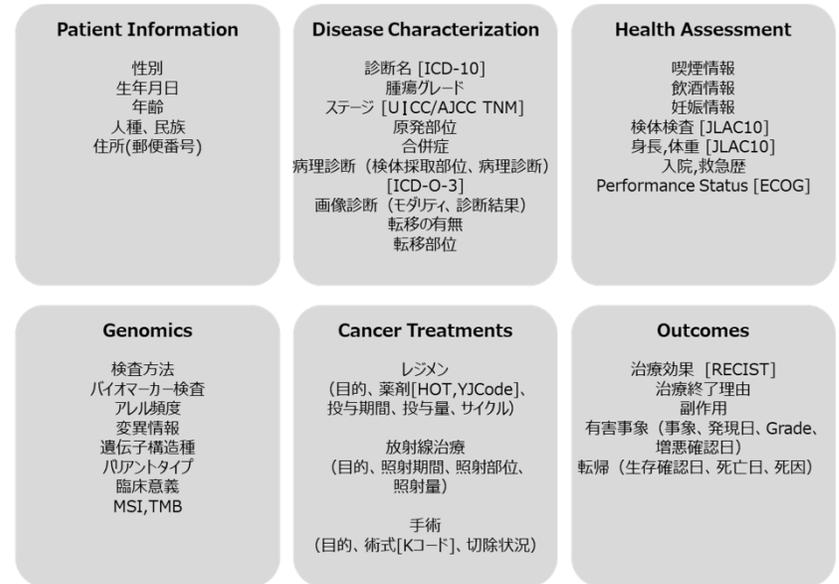
HL7 FHIRに準拠した多施設連携可能な診療情報基盤の構築

▶ HL7 FHIRに準拠した多施設間連携が可能な診療情報収集基盤

- ✓ 基幹電子カルテシステムとクラウドサービスを用いたデジタル医療データバンク
- ✓ 電子カルテや部門システムに散在するがん領域データセットをデータバンクに集約
電子カルテ、放射線、放射線治療、病理検査、医事会計、画像システム、からのデータ連携と利活用
- ✓ SIP3-D1（東大・河添班）に準拠したデータ形式
- ✓ SIP3-B1（がん研有明）と連携する収集データ項目の検討(意見交換会の定例開催)



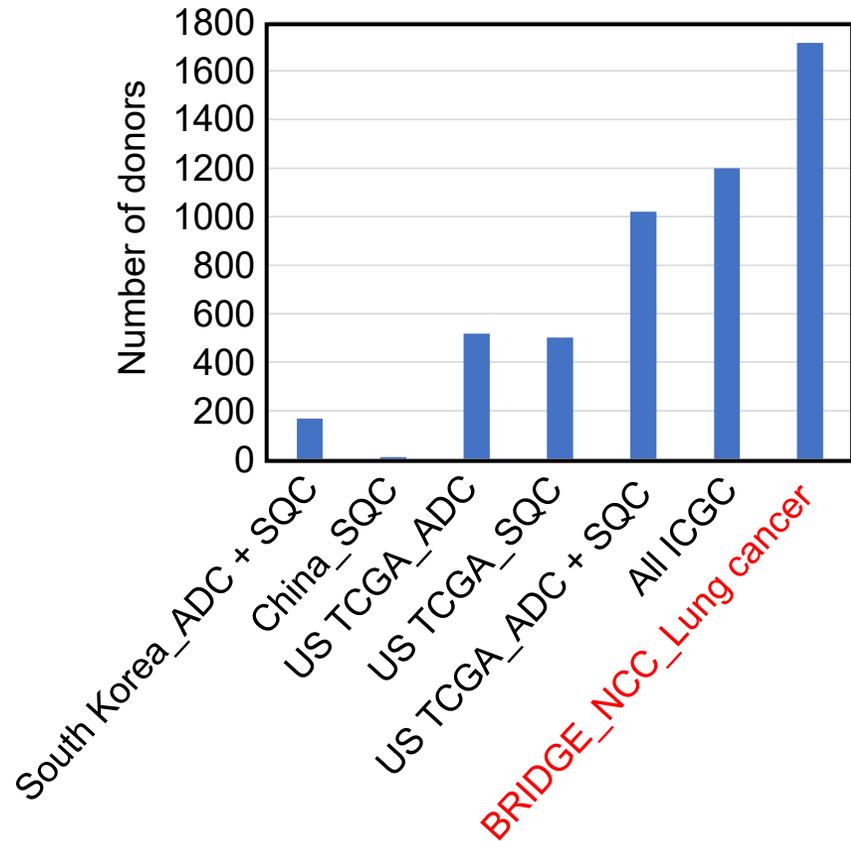
多施設間でのデータ収集に対応した診療情報収集基盤



がん領域における収集データ項目例

AI解析を志向した世界最大規模の肺がん多層オミックスデータベースの構築

肺がんデータベースにおける症例数の比較



肺がん統合データベースの内訳

	TCGA		BRIDGE + がんゲノムプロジェクト
	ADC	ADC + SQC	
臨床情報 オミックスデータ	585	1,089	1,913 + 2,294 = 4,207
全エクソーム解析	582	1,084	1,965
RNA-seq解析	515	1,016	1,809 + 754 = 2,563
DNAメチル化解析	579	1,082	403*
全ゲノム解析	99	-	462* + 878 = 1,291
ChIP-seq (H3K27Ac)解析	-	-	222*

*全ゲノム解析・ChIP-seq・DNAメチル化解析に関しては、pan-negative肺がん症例に集中して解析を行った。

ICGC Data Portal_lung cancer (<https://dcc.icgc.org/q?q=lung%20cancer>) に2023年3月8日にアクセス

2023年以降に発表した主要な論文:

Nature Communications 14:4991 (2023), *Nature Communications* 14:3043 (2023), *British Journal of Cancer* 129:2003-2013 (2023), *Journal of Thoracic Oncology* 19:984-994 (2024), *Cancer Communications* 44:287-293 (2024), *Experimental Molecular Medicine* 56:646-655 (2024), *Molecular Cancer* 23:126 (2024), *Molecular Cancer* 23:182 (2024), *npj Systems Biology and Applications* 11:49 (2025), *Journal of Thoracic Oncology* S1556-0864(25)00764-6 (2025)

医用画像データベースの構築の例（内視鏡画像バンク）

内視鏡画像バンクの構築（動画・静止画）

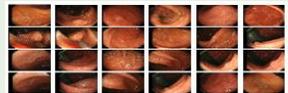
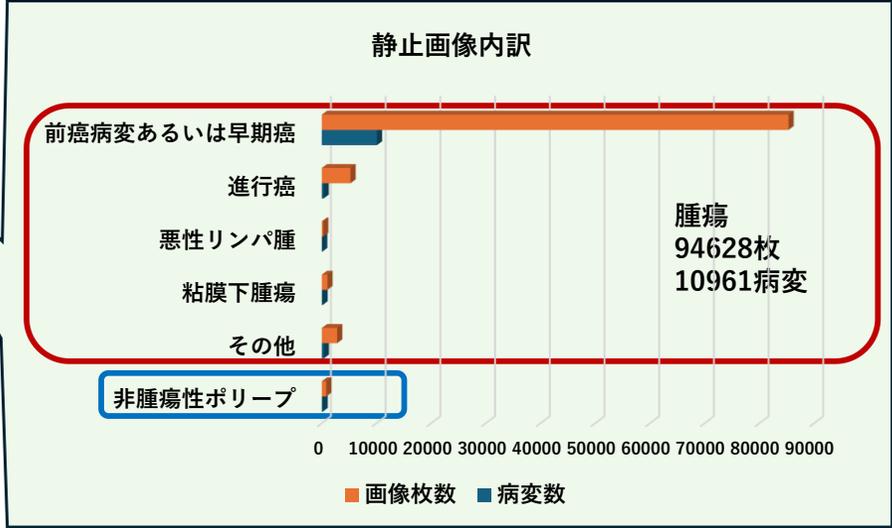
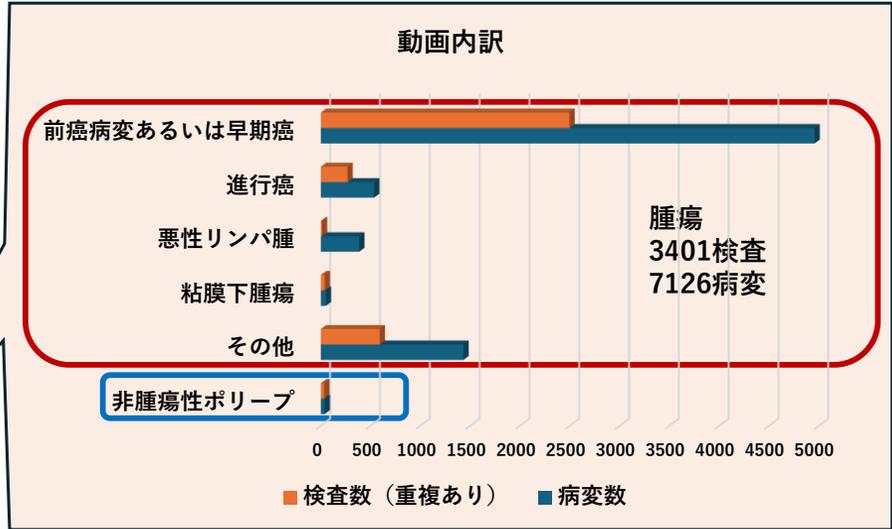
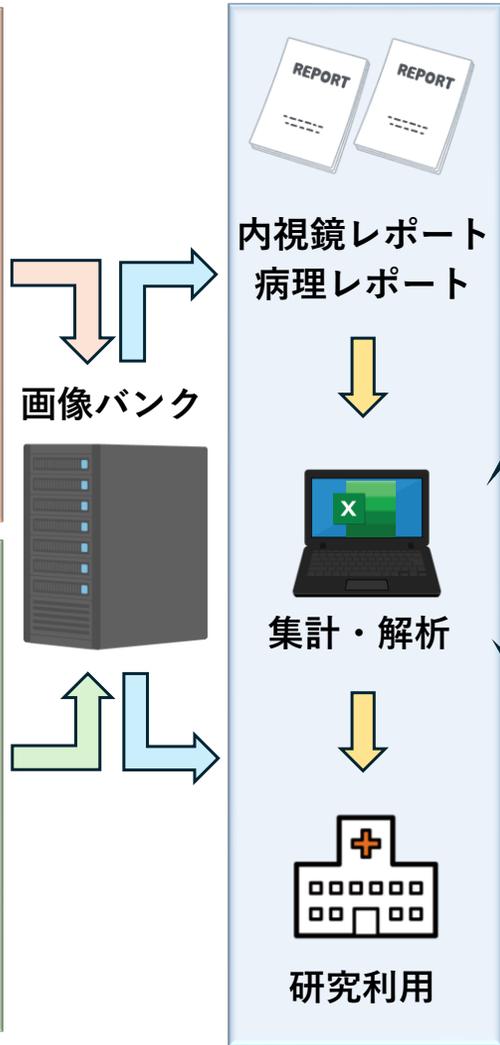
大腸内視鏡検査動画

検査年	検査数
2017	611
2018	1852
2019	90
総計	2553



大腸内視鏡検査静止画像

検査年	病変数	画像数
2013	1498	8667
2014	531	4920
2015	462	5218
2016	516	4609
2017	4595	17343
2018	2426	20207
2019	1077	34474
総計	11105	95438

医用画像データベースの構築の例 (超音波画像バンク・悪性脳腫瘍における発散的拡散強調画像バンク)

超音波画像バンクの構築 (動画・静止画)

▶ 胎児心臓超音波検査 (先天性心疾患)

		Case	Video	B-mode	Color flow	equipment
妊娠中期	正常例	770	2054	1653	401	3社9機種
	異常例	64	359	266	93	2社4機種
妊娠初期	正常例	185	780	409	371	1社2機種
	異常例	15	187	51	136	1社4機種

▶ 経膈超音波検査 (子宮内膜癌)

	Case	Image
正常例	117	165
異常例	74	134
合計	191	299

▶ 内視鏡超音波検査 (膵腫瘍)

	Case	Video
異常例	60	60

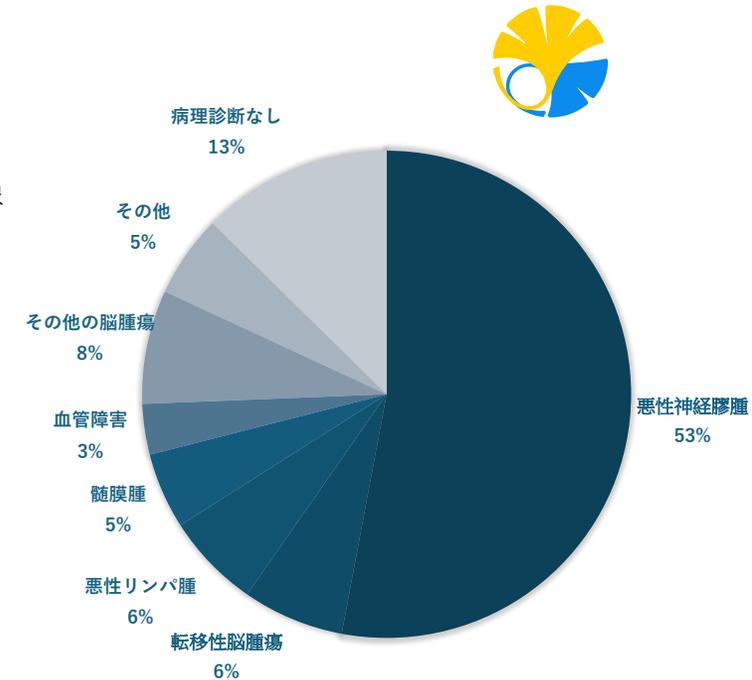
▶ 経気管支超音波検査 (肺癌)

	Case	B-mode	Color flow	Elastography
異常例	204	452	452	453

悪性脳腫瘍における発散的拡散強調画像バンク

東京大学医学部附属病院にて
全症例 1社1機種のみMRI装置で撮像
症例総数: 238

悪性神経膠腫: 126,
転移性脳腫瘍: 16,
悪性リンパ腫: 16,
髄膜腫: 12,
血管障害: 8,
他の脳腫瘍: 18,
その他脳疾患: 13,
病理なし: 29



国立がん研究センター研究所医療AI研究開発分野がこれまで 社会実装/臨床応用に貢献した成果（実証実験を含む）

製品化

1. 内視鏡診断支援AIシステム（大腸病変検知）【日本・薬事承認：管理医療機器（Class II）】
商品名：WISE VISION（NEC）2020年11月
2. 内視鏡診断支援AIシステム（大腸病変検知）【欧州・CEマーク取得】
商品名：WISE VISION（NEC）2020年12月
3. 内視鏡診断支援AIシステム（バレット食道の腫瘍検知）【欧州・CEマーク取得】
商品名：WISE VISION（NEC）2021年5月
4. AI開発支援プラットフォーム【上市】
商品名：SYNAPSE Creative Space（FUJIFILM）2022年4月
5. MRI画像から神経膠腫の疑いのある領域を精密に抽出するAI技術【日本・認証：管理医療機器（Class II）】
商品名：SYNAPSE Creative Space（FUJIFILM）2024年2月
6. 胎児心臓超音波診断支援AIシステム【日本・薬事承認：管理医療機器（Class II）】
商品名：HOPE LifeMark 胎児心臓超音波スクリーニング支援システム（富士通）2024年7月
7. 内視鏡診断支援AIシステム（大腸鑑別診断）【日本・薬事承認：高度管理医療機器（Class III）】
商品名：WISE VISION 内視鏡画像解析AI A1000（NEC）2025年3月

自治体との共同事業（実証実験）

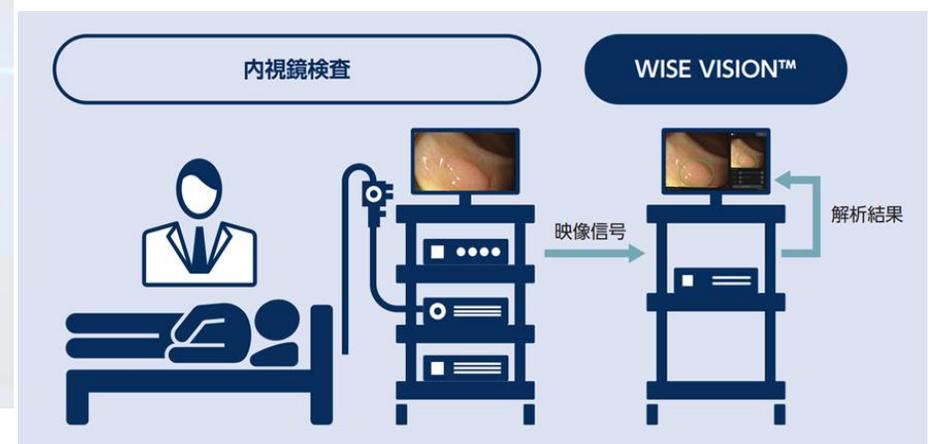
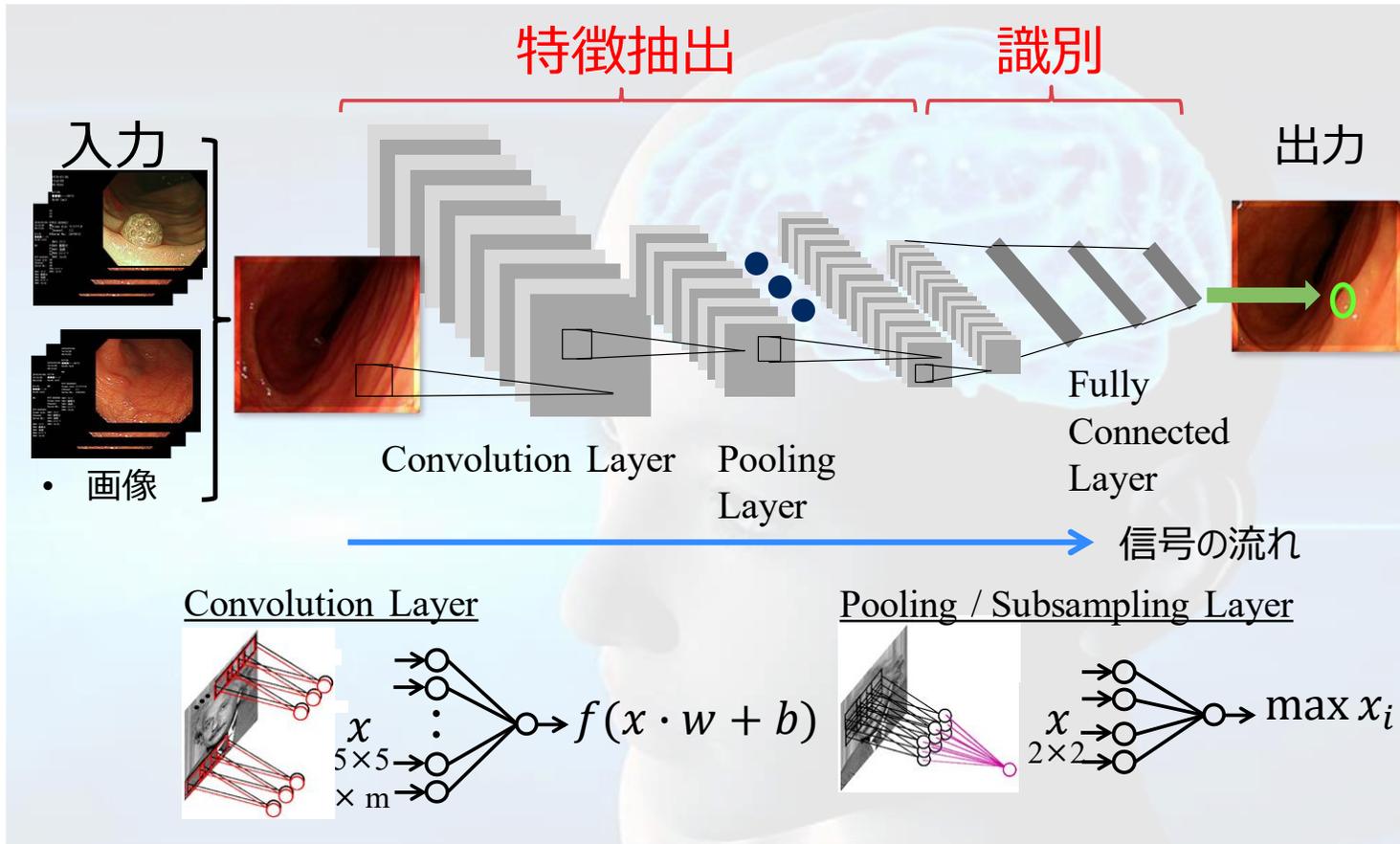
AI及びリモートテクノロジーを用いた心房細動の早期発見により清水区を脳梗塞の少ないまちにする地域医療プロジェクト(略称SPAFS)：
静岡市清水区・東京医科歯科大学・フクダ電子等との共同事業

→2023年1月末時点で解析が終了していた336人（総参加者350人）中、9人（およそ2.7%）に心房細動が見つかった。

成果例：AIを活用したプログラム医療機器 (SaMD) の薬事承認取得

WISE VISION™：内視鏡用AI診断医療機器の開発

大腸がんおよび前がん病変発見のためのリアルタイム内視鏡診断サポートシステムに関して2020年に管理医療機器 (Class II) として薬事承認を受け (承認番号：30200BZX00382000)、また欧州においても医療機器製品の基準となるCEマークの要件に適合した。また、バレット食道の腫瘍検知にも応用し、世界で初めて製品としてCEマークに適合した。



成果例：超音波診断支援AIの実臨床応用（薬事承認の取得）

超音波診断支援AIの実臨床応用

－胎児心臓超音波スクリーニング支援システムの薬事承認－

📄 ポスト 📄 シェアする 📄 LINEで送る 1K

2024年9月6日

国立研究開発法人理化学研究所
国立研究開発法人国立がん研究センター
昭和大学

概要

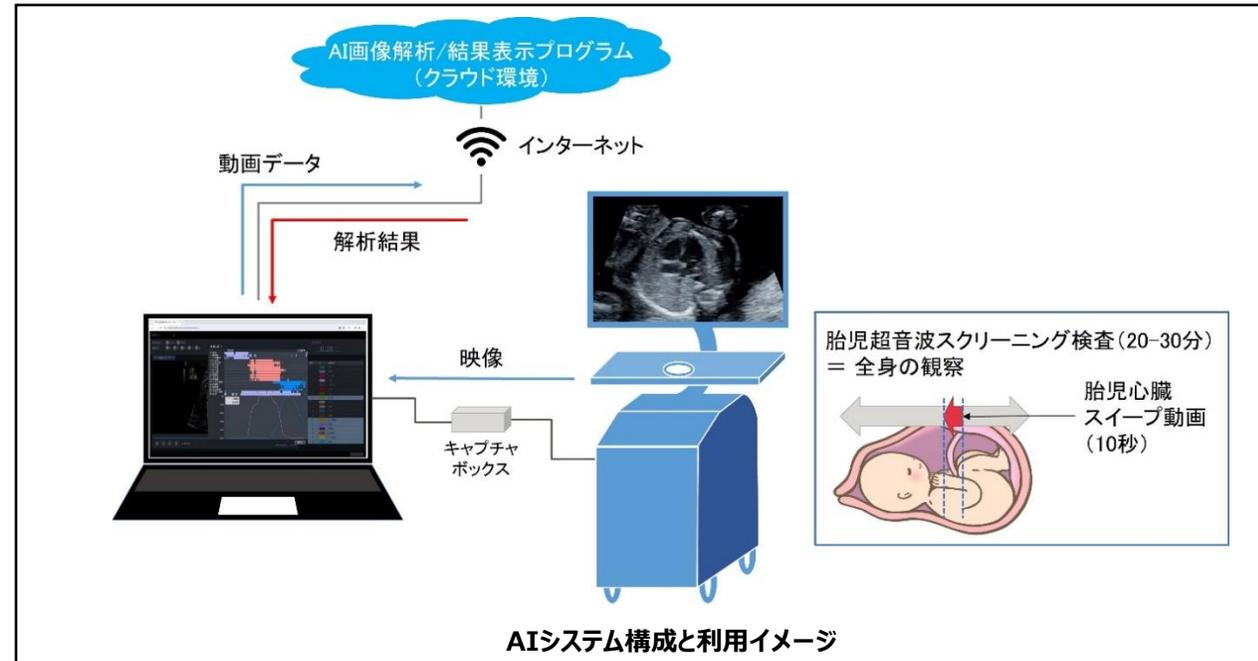
理化学研究所（理研）革新知能統合研究センター目的指向基盤技術研究グループがん探索医療研究チームの小松正明副チームリーダー、国立がん研究センター研究所医療AI研究開発分野の浜本隆二分野長（理研革新知能統合研究センター目的指向基盤技術研究グループがん探索医療研究チーム・チームリーダー）、昭和大学医学部産婦人科学講座の松岡隆准教授らの共同研究グループは、超音波検査で胎児の心臓に異常がないかスクリーニングする際に、検査者の超音波診断を支援する人工知能（AI）システムを共同開発し、AI搭載医療機器プログラムとして厚生労働省の薬事承認を取得しました（2024年7月29日付）。

本研究成果は、胎児の超音波診断を支援するとともに、早急に治療が必要な重症かつ複雑な先天性心疾患の見落としを防ぎ、早期診断や綿密な治療計画の立案につながると期待できます。また、日本の少子化および産婦人科医の不足や偏在など切迫した状況において、検査者間の技術格差や地域間の医療格差を埋めることで、産婦人科医療のさらなる発展に貢献するものと考えられます。

本研究開発は、国立がん研究センター（理事長：中釜育）研究所と中央病院、理研革新知能統合研究センターが連携して取り組んでいる内閣府BRIDGE事業「医療デジタルツインの発展に資するデジタル医療データバンク構想」におけるプロジェクトの一つとして実施されました。また本研究開発において、昭和大学は通常の診療にて妊婦（胎児）の超音波画像情報を取得し、国立がん研究センターにてデータ収集・統合を行いました。次に、その統合したデータを理研にて解析し、新規アルゴリズムを獲得しました。さらに、国立がん研究センターと昭和大学においてアルゴリズムの検証やプログラムの性能評価を行いました。

出典：国立がん研究センタープレスリリース

(https://www.ncc.go.jp/jp/information/pr_release/2024/0906/index.html) 2024年10月11日アクセス



将来の目標

1. 出生前検査としてルーチンで行われる、**胎児超音波スクリーニングへの標準装備を目指す。**
2. 検査者のトレーニングやクラウド化による遠隔診断の実現で、地域間の医療格差の大幅な是正が行われ、医師不足に悩む周産期、新生児医療の発展を通して、**さらなる安全な妊娠・出産を実現させる。**

成果例：マルチオミクス解析により、肺腺がんの分子メカニズムを解明 —有効な治療薬のなかった肺腺がんの創薬に向けた研究に貢献—

マルチオミクス解析により、肺腺がんの分子メカニズムを解明
有効な治療薬のなかった肺腺がんの創薬に向けた研究に貢献

📄 ポスト 🌐 シェアする 📞 LINEで送る 👤 1K

2024年9月13日

国立研究開発法人国立がん研究センター
国立研究開発法人理化学研究所

発表のポイント

- 肺腺がんのサブタイプ、なかでも、分子標的治療の対象外である症例について、全ゲノムデータ、エピゲノムデータ（クロマチン免疫沈降解析シーケンスおよびDNAメチル化）、トランスクリプトームデータを活用して、統合解析を行いました。
- 治療標的となる遺伝子変異（ドライバー変異）が見つからない肺腺がん症例において、MAML2遺伝子の発現が有意に低下していることが明らかとなりました。
- MAML2遺伝子との相関解析から、患者予後予測に関連するマーカー遺伝子を同定しました。
- 今回の発見は、肺腺がんの新たな予後マーカーとして創薬に応用できる可能性を示唆するとともに、機械学習を用いることでより精緻に患者を層別化できることを示しました。

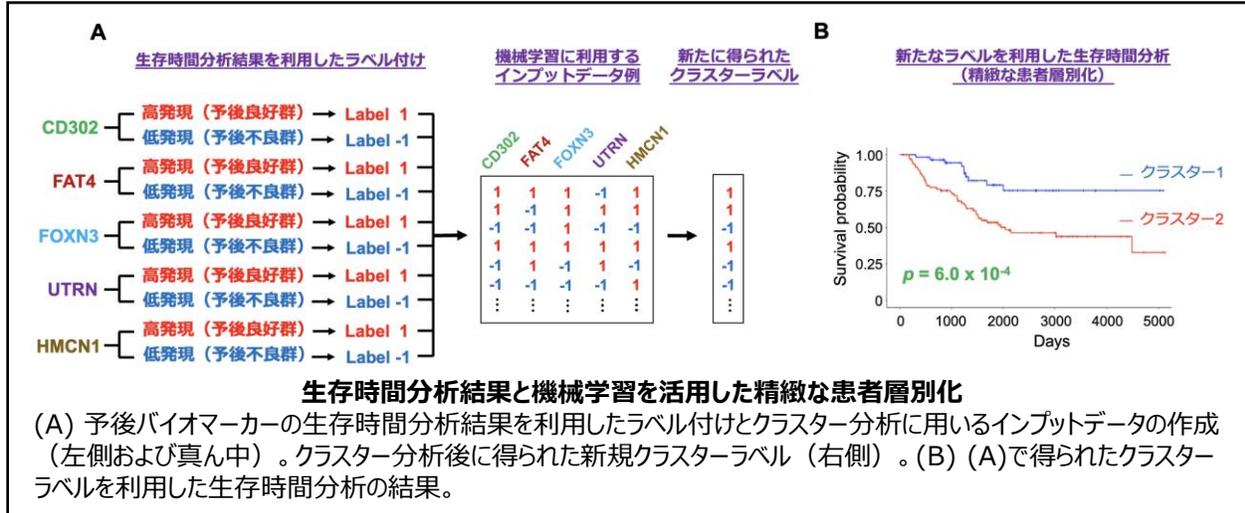
【掲載雑誌】

Molecular Cancer 23:182 (2024)

IF = 33.9 (Journal Impact Factor - Journal Citation Reports 2024)

出典：国立がん研究センタープレスリリース

(https://www.ncc.go.jp/jp/information/pr_release/2024/0913/index.html) 2024年10月11日アクセス



生存時間分析結果と機械学習を活用した精緻な患者層別化

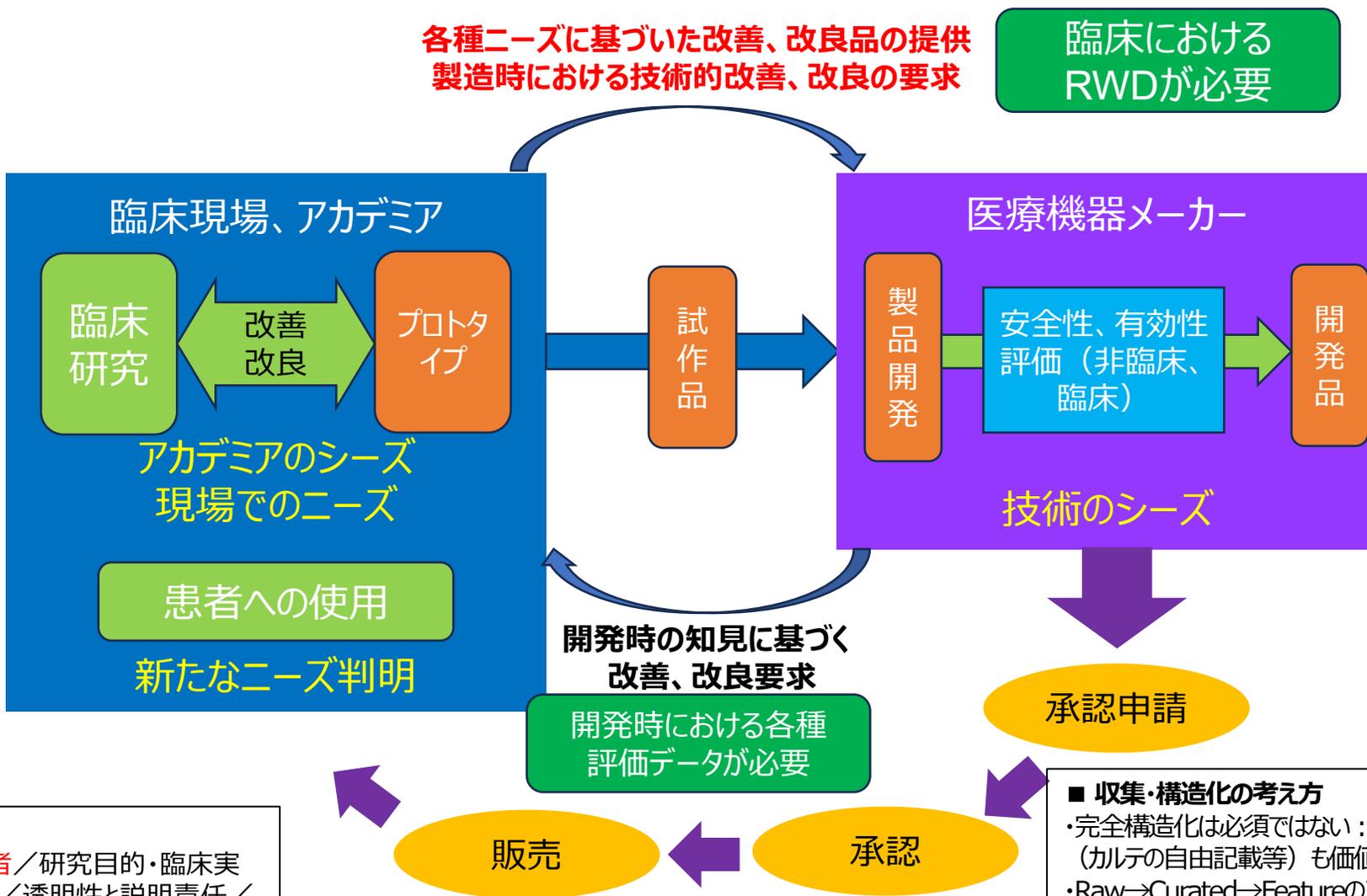
(A) 予後バイオマーカーの生存時間分析結果を利用したラベル付けとクラスター分析に用いるインプットデータの作成（左側および真ん中）。クラスター分析後に得られた新規クラスターラベル（右側）。(B) (A)で得られたクラスターラベルを利用した生存時間分析の結果。

本研究では、分子標的治療の適応外となる肺腺がん症例を対象に、全ゲノム、エピゲノム（ChIP-seqおよびDNAメチル化）、トランスクリプトームの各データを統合解析した。その結果、発がんに関与する新規の分子機構を明らかにし、創薬へとつながる可能性を見出した。さらに、AI（機械学習）を用いた解析により、患者の精緻な層別化が可能であることも示した。これらの成果は、肺がん研究の進展に寄与するとともに、新たな創薬標的の探索基盤となることが期待される。

医療AI研究開発においてゲノム情報を中心としたオミクス情報は重要である。

医療機器の開発におけるRWDの重要性

【一般的な医療機器の開発イメージ】



医療機器は改善・改良の繰り返りで優れた製品、現場に必要な製品が生まれる



SaMDに限らず、医療機器の改善・改良には臨床におけるRWDが非常に重要（業界団体も指摘）

RWDの利活用により

- ・改善・改良の推進（PDCAサイクルの加速）
- ・治験数の削減
- ・本来困難な医療貢献度等の数値化の可能性

等が期待される。

■ 基本原則

患者さんは協働者／研究目的・臨床実装目的を明文化／透明性と説明責任／法令倫理の順守／データ来歴と意思決定の説明可能性を担保

■ 収集・構造化の考え方

- ・完全構造化は必須ではない：非構造データ（カルテの自由記載等）も価値あり
- ・Raw→Curated→Featureの三層アーキテクチャ：メタデータは必ず構造化して追跡
- ・NLP/画像アノテーションで“非構造→半構造”を段階的に

AI搭載医療機器を開発するうえでの課題（開発側）

1. 医用画像データの公共整備と質保証（次世代医療基盤法の実効化）

企業が利用可能な公共の医用画像（X線・CT・MRI ほか）の整備が次世代医療基盤法では依然遅れており、他検査との連携によるマルチオミクス診断の推進が十分でない。数量の確保だけでなく、質（正解情報＝アノテーション）の担保が不可欠であり、医師の関与なしには実現しないが、アノテーション負荷への補償・支援が不足している。

2. 学習用データの継続利用に関する理解促進

AI開発では、学習用データを永続的に使用する必要がある（例えば、共同研究契約などで入手し製品化に使用したデータについて、研究契約期間が終了後もバージョンアップ時の機能確認などで過去に使用したデータは利用必須である）。このことが医療機関側で理解されないケースがあり、開発に支障を来たすことがある。このような状況について、医療機関や認可に関連する組織への啓発活動が欲しい。

3. 越境データ利活用の枠組み

外資系に限らず国内企業でも海外に開発拠点を置くケースが多い。収集データのグローバル相互利用（W/W展開）を可能にする、法令適合的で実務に沿った枠組みが必要ではないか。

4. 仮名加工情報の共同利用に関する運用明確化

個人情報保護法における仮名加工情報の共同利用について、医療機関向けガイドラインでは「オプトアウトにより企業が商品化にも利用可能」と記載があるものの、運用実績が乏しい。データ活用を強く求めるAI分野では時間的猶予が小さいため、企業との共同利用スキームや手続の明確な記載・具体的運用指針の早期提示を期待する。

5. 次世代医療基盤法データの実効的活用

次世代医療基盤法の仮名加工医療情報を企業が認定利用事業者として活用する場合、今後多くの企業が必要とする、疾患特化を含む多様な医用画像・病理データセットを、グローバル競争に耐える質・量で、容易かつ迅速に整備・入手できるのが課題である。過去のデータ整備の経緯からも困難が予想されるため、標準化・品質保証・流通促進のための具体策（例：共通仕様、収集・匿名化標準、提供期限・SLA、料金・権利整理の明確化）の検討・実装が必要ではないか。

6. 日本語学術コンテンツのデータ化・利活用の加速

海外と比べ、AI学習に直接利用できる日本語コーパスが著しく不足している。紙面・PDFの学術文献は多いが、機械可読な形へのデータ化・権利処理・メタデータ付与が追いついていない。生成AIの拡大に伴い、学習資源の枯渇が喫緊の課題となるため、早急な対応が必要ではないか。

7. AI搭載医療機器の制度設計

従来の薬機法の範囲を背景として規定・規制するのではなく、医師の利用方法などを考慮して、医師法/医療法を背景とする新しい法規制の枠組みの策定や市場展開の方法など、（AI関連技術の開発スピードに見合った）医療機器としての迅速な製品化や市場導入を推進するための支援策を検討して欲しい。

補足資料

成果例：MRI画像から神経膠腫の疑いのある領域を精密に抽出するAI技術を共同開発（薬機法下における認証の取得）

富士フィルムと国立がん研究センター
MRI画像から神経膠腫の疑いのある領域を精密に抽出するAI技術を共同開発
希少がんである神経膠腫の治療前の画像評価精度を向上

✕ ポスト シェアする LINEで送る 1K

2024年2月28日
富士フィルム株式会社
国立研究開発法人国立がん研究センター

概要

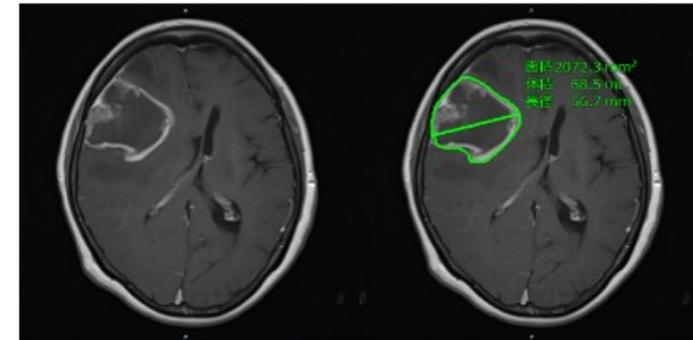
富士フィルム株式会社（本社：東京都港区、代表取締役社長・CEO：後藤 禎一）と国立研究開発法人国立がん研究センター（所在地：東京都中央区、理事長：中釜 斉）は、MRI画像から神経膠腫（グリオーマ）の疑いのある領域を精密に抽出するAI技術^{*1}を共同で開発しました。本技術により、希少がんである神経膠腫の治療前の画像評価の精度向上が期待できます。

本技術は、富士フィルムのAI技術開発支援サービス「SYNAPSE Creative Space（シナプスクリエイティブスペース）」を活用して開発したAI技術です。今後、富士フィルムは、本サービスを活用して開発したAI技術の初めての社会実装に向けて、本技術を搭載した製品の早期市場導入を目指します。

出典：国立がん研究センタープレスリリース
（https://www.ncc.go.jp/jp/information/pr_release/2024/0228/index.html） 2024年10月11日アクセス

研究成果

今回、富士フィルムと国立がん研究センターは、国立がん研究センター研究所 医療AI研究開発分野の小林 和馬 研究員および浜本 隆二 分野長、同中央病院・脳脊髄腫瘍科の高橋 雅道 医長、同中央病院・放射線診断科の三宅 基隆 医長らを中心としたチームのもと、MRI画像から神経膠腫の疑いのある領域を精密に抽出するAI技術を開発しました。本技術は、富士フィルムの「SYNAPSE Creative Space」により、頭部MRI画像から神経膠腫の領域を抽出するアノテーション作業を効率的に行い作成したデータをAIに学習させて開発したものです。MRI画像から神経膠腫の疑いのある領域を精密に抽出し、抽出した領域の体積を計測できます。本技術により、神経膠腫の治療前の画像評価を、より高い精度で行えることで、将来的に、早期発見や診断の精度向上、放射線治療や手術などの治療計画の最適化などに役立つことが期待できます。



神経膠腫の疑いのある領域の抽出結果の例

成果例：受動喫煙が肺がんの遺伝子変異を誘発することを証明

受動喫煙が肺がんの遺伝子変異を誘発することを証明

受動喫煙を回避することの重要性を肺がんの発症機構からも確認

× ポスト

シェアする

LINEで送る

1K

2024年4月16日

国立研究開発法人国立がん研究センター
国立大学法人東京医科歯科大学

発表のポイント

- 受動喫煙は、肺がんの危険因子として知られていますが、受動喫煙と遺伝子変異との関わりは不明でした。
- 肺がん女性の受動喫煙歴と遺伝子変異の関係を調べた結果、受動喫煙を受けて発生した肺がんでは、受動喫煙を受けずに発生した肺がんと比べて、より多くの遺伝子変異が蓄積していることがわかりました。
- 受動喫煙は、能動喫煙とは異なるメカニズムで変異を誘発し、肺の中にできた初期の腫瘍細胞が悪性化するのを促進すると推定されました。
- これらの発見は、受動喫煙による健康被害を防ぐ必要性を強く示唆しており、また受動喫煙による肺がんの予防に役立つことが期待されます。

【掲載雑誌】

Journal of Thoracic Oncology 19:984-994 (2024)

IF = 20.8 (Journal Impact Factor - Journal Citation Reports 2024)

出典：国立がん研究センタープレスリリース

(https://www.ncc.go.jp/jp/information/pr_release/2024/0416/index.html) 2024年10月11日アクセス

肺がんのゲノム解析

	非喫煙者女性		能動喫煙者女性
	受動喫煙なし	受動喫煙あり	
100万塩基中の遺伝子変異数*	1.29	1.44	1.99
APOBEC型変異の割合 (%)*	7.32	15.6	0
たばこ型変異の割合 (%)*	0	0	33.5

*中央値を示す

受動喫煙は能動喫煙と違う変異を誘発する

10歳代、30歳代のいずれか、あるいは両方で受動喫煙を受けていた（月に1～2日から毎日まで）方に生じた肺がんでは、受動喫煙を受けていない方の肺がんと比べて、より多くの遺伝子変異が蓄積していました。また、能動喫煙者の肺がんで見られるたばこ中の発がん物質により直接引き起こされるタイプの変異が、受動喫煙者の肺がんではごく稀にしかみられませんでした。よって、受動喫煙は能動喫煙とは違うメカニズムで変異を誘発することが明らかになりました。

AI/ML-SaMDにおける課題

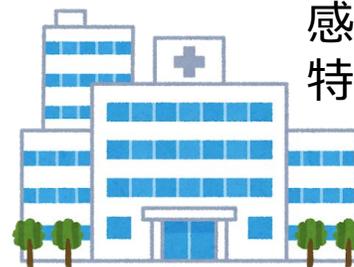
開発時と画質などの異なるデータに対して使用した場合に、**Real World (RW)** で期待通りの性能が得られないことがある

<例>

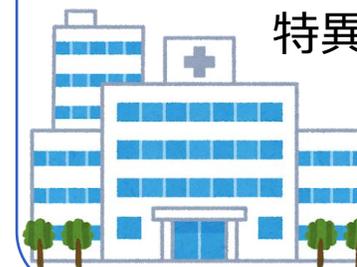


感度90%、
特異度95%
を確認

承認
販売



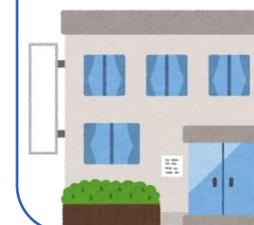
感度92%、
特異度95%



感度70%、
特異度75%



感度95%、
特異度85%



感度70%、
特異度95%

開発コンセプト、学習及びテストデータなどの実態や分布がRWと一致していないことが主な要因（ドメインシフトの発生）

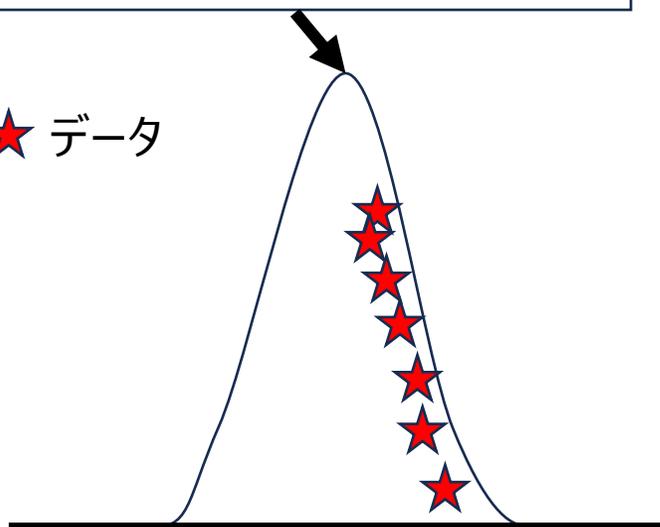
RWとの性能ギャップ

<主要要因>

- 学習データ、性能評価データ量が母集団を十分に反映していない
→十分な数のデータを使用できない場合、あるいは使用したデータが母集団の代表となっていない場合、ML利用SaMDの性能はRWで低下する可能性がある

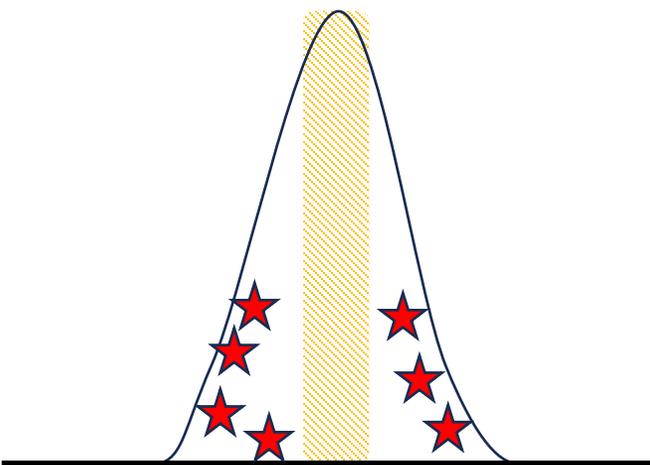
(例) **RWにおける対象疾患の分布**

★ データ

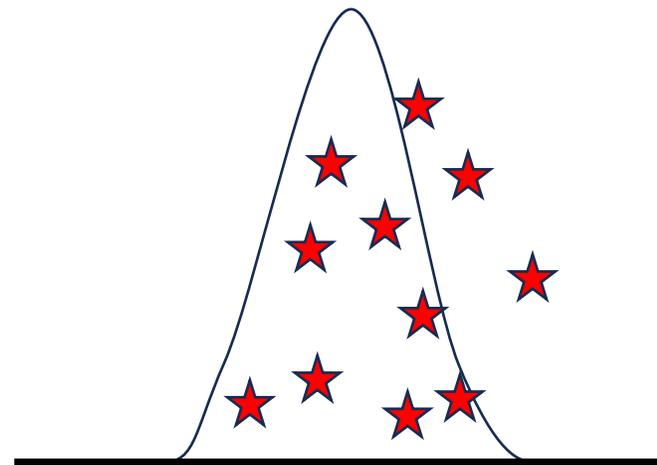


利用データの
特定症状への偏り

収集したデータを学習、性能評価のいずれに使用しても得られたSaMDの性能はRWと乖離が生じる



利用データにおける
特定症状の不足

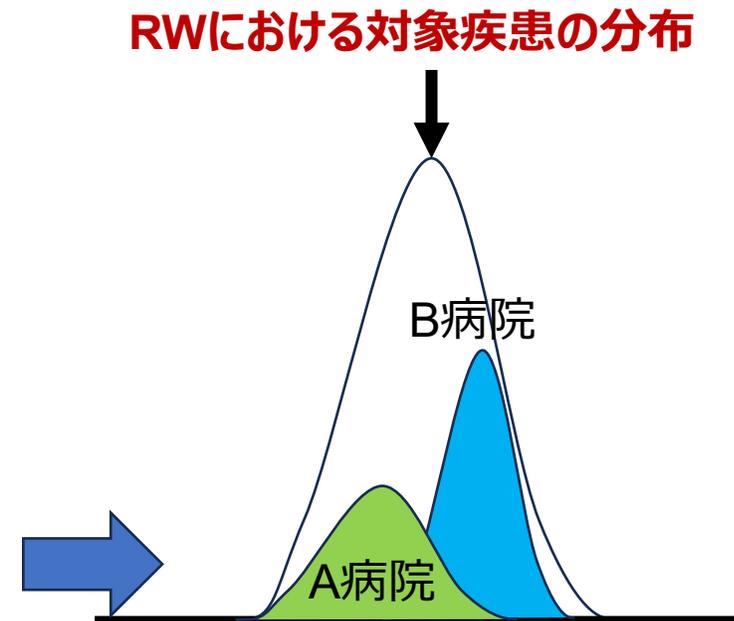


利用データにおける
例外的症状の存在

偏りが生じないほどの数を事前に収集することは現実的に困難→市販後のRWD収集で補えないか？

<その他の要因（RWDに影響する要因）>

- 使用装置や撮影条件の違い（施設間の差異）
 - 装置全般：メーカ、製品、バージョン違い
 - CT：スキャン方法、マルチスライス数、再構成フィルタ等
 - MR：磁場強度、RFコイルの種類、シーケンス、撮影時間等
 - 内視鏡：照明条件、コントラスト等
- 対象母集団分布の違い（対象患者群の差異）
 - 受診する患者の年齢、性別、疾患の進行度等が病院毎に異なる等



要因 = 考慮すべき課題

- ・利用可能なデータ数の少なさ（必要なデータの収集も未だ困難）
- ・データに付随する情報と要求される情報との不一致（匿名化により学習に必要な情報が削除されうる）
- ・仮名加工情報のメリット、デメリットが不明瞭なため、活用が不十分（ユーザーの不勉強もあるが、もう少し分かり易い仕組みとすべき）
- ・有効性等を担保する上で必要な、科学的に妥当な各種データ数が不明瞭

課題解決にはRWDの利活用を進めることが有効な手段の一つと考えられる

⇒SaMDに限らず医療向上に資するRWD利活用推進の新たな枠組みが必要では？

資料作成にご協力いただいた先生方（五十音順）



- ・浅田 健 先生
- ・金子 修三 先生
- ・河野 隆志 先生
- ・河野 伸次 先生
- ・小林 和馬 先生
- ・小松 正明 先生
- ・斎藤 豊 先生
- ・白石 航也 先生
- ・高橋 慧 先生
- ・田中 勝弥 先生
- ・山田 真善 先生

- ・板倉 陽一郎 先生
- ・井元 清哉 先生
- ・鎌谷 洋一郎 先生
- ・瀬々 潤 先生
- ・田宮 元 先生
- ・中岡 竜介 先生
- ・舟橋 毅 先生
- ・山口 類 先生
- ・山本 陽一郎 先生