



国際医療福祉大学
INTERNATIONAL UNIVERSITY OF HEALTH AND WELFARE



cardiologist, specialist for pulmonary hypertension, cardio-oncology and AI digital health. 循環器専門医(=心臓専門)。肺高血圧症・腫瘍循環器・AI/デジタルヘルスを専門にしています

日本脳卒中学会/循環器学会 脳卒中と循環器病克服 第2次5カ年計画委員



市販前の新たな治験実施の有無によらず、プログラム医療機器を含む医療機器の承認試験を行い得ると考えられるケースの取扱いについては、『医療機器の臨床試験の試験取組に関する資料』の提出が必要範囲等に係る取扱い（市販前・市販後を通じた取組みを踏まえた対応）について（薬生審査書第1117第1号、薬生安第1117第1号、平成29年11月17日厚生労働省医薬・生活衛生局医療機器審査管理課長、医薬・生活衛生局医薬安全対策課長連名通知。以下、「リバランス通知」という。）により示しています。

今般、プログラム医療機器の開発及び薬事承認を効率的に行う観点から、「プログラム医療機器の特性を踏まえた薬事承認制度の運用改善検討事業」(研究代表者: 佐藤正樹、(株)日本医療機器開発センター、産業医事局)において、プログラム医療器

医療機器・IVD工業会、欧州ビジネス協会医療機器・IVD委員会、日本デジタルヘルス・アライアンス、一般社団法人日本医療ベンチャー協会、A I 医療機器協議会、日本製薬工業協会及び独立行政法人医薬品医療機器総合機構宛てに送付していただくことを申し上げます。



厚生労働省では医療ニーズが高く実用可能性のある次世代医療機器について、審査時に用いる技術評価指標等をあらかじめ作成し公表することにより、製品開発の効率化及び承認審査の迅速化を図る目的で、検討分野を選定して評価指標を公表してきたところです。

今般、乳がん診断支援装置及び行動変容を伴う医療機器プログラムの評価を行うに当たって必要と考えられる資料、評価のポイント等を別紙1～2のとおり評価指標としてとりまとめましたので、下記に留意の上、製造販売承認申請に際してお考えとす

治療用アプリの承認基準策定



**2050年の医療/医療技術の未来図白書
『医療技術フォーサイト 2050』を
2025年7月に公開
(公益財団法人 医療機器センター)**

AIの進化と医療・医療機器への関与

医療現場の業務支援
(AIエージェント=非医療機器)

診断・治療支援
(医療機器)

AI-driven
診断・治療
(デバイス一体化医療機器)

因果律を超えた
アウトカムベースの
AI機器？

退院サマリー自動生成システム
AI音声外来予約管理システム



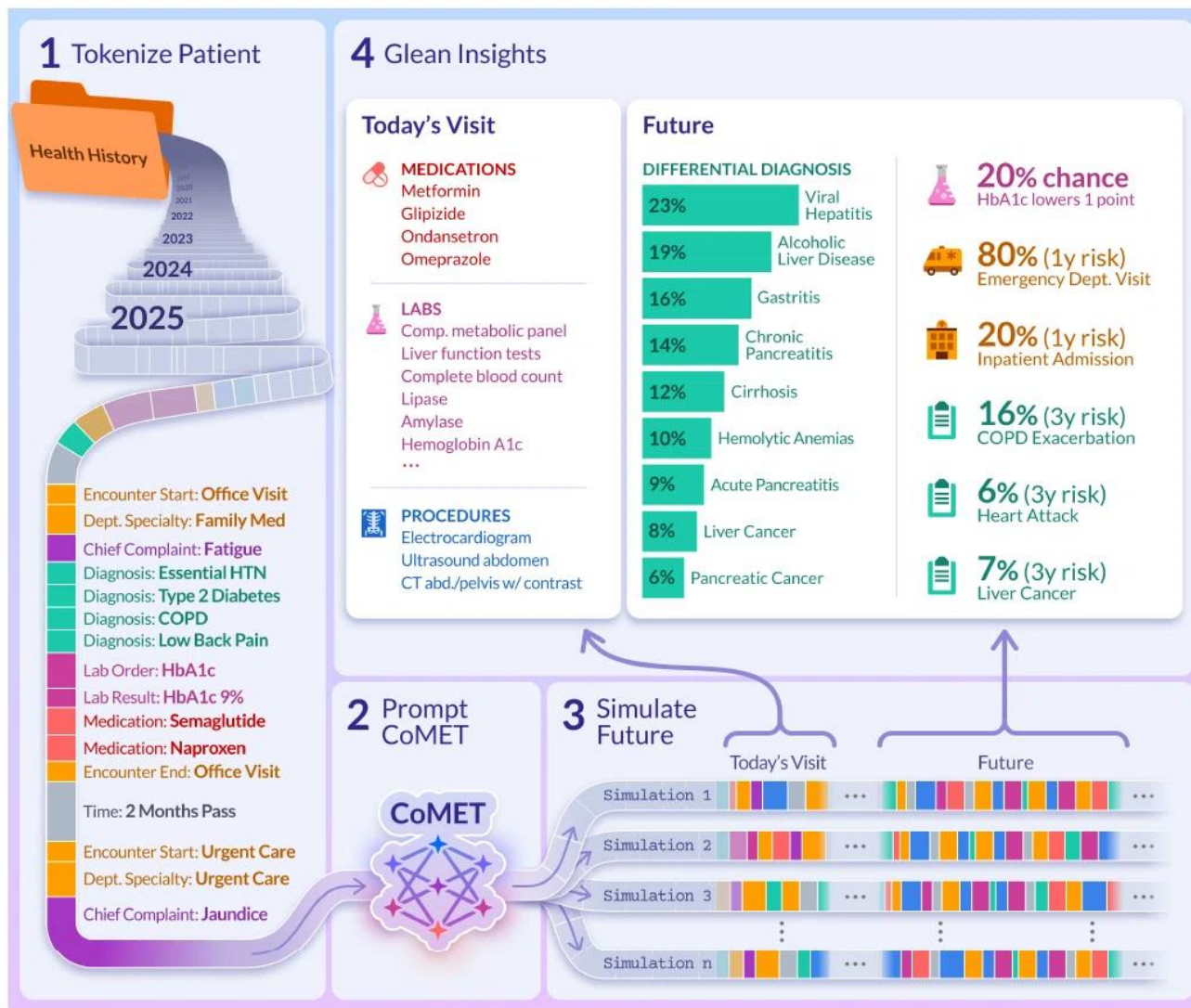
長時間心電図AI解析システム
AIインフルエンザ検査



疾患兆候/治療反応性検出
クローズドループ人工臓



イベントを予測するAI研究



- 3億人の電子カルテ情報から次の臨床イベントを予測するシステム

慢性疾患の増悪
救急受診リスク

など

検診データから実年齢と臓器年齢の乖離を評価?

nature medicine

Article

<https://doi.org/10.1038/s41591-025-03856-8>

Large language model-based biological age prediction in large-scale populations

Received: 1 December 2024

Accepted: 20 June 2025

Published online: 23 July 2025

Check for updates

A list of authors and their affiliations appears at the end of the paper

Accurate and convenient assessment of individual aging is crucial for identifying health risks and preventing aging-related diseases. Nonetheless, current aging proxies often face challenges such as methodological limitations, weak associations with adverse outcomes and limited generalizability. Here we propose a framework that leverages large language models (LLMs) to estimate individual overall and organ-specific aging using only health examination reports. We validated this approach across six population-based cohorts, encompassing over 10 million participants and demonstrated effectiveness and reliability. Our results showed that the LLM-predicted overall age achieved a concordance index (C-index) of 0.757 (95% CI 0.752–0.761) for all-cause mortality, significantly outperforming other aging proxies such as telomere length, frailty index, eight epigenetic ages and four machine-learning models predictions. The overall age gap was strongly associated with multiple aging-related phenotypes and health outcomes, showing a hazard ratio of 1.055 (95% CI 1.050–1.060) for all-cause mortality. For organ-specific aging, LLM-predicted ages and age gaps also demonstrated superior performance in predicting corresponding organ-specific diseases compared to machine-learning models. Additionally, we examined the dynamic aging assessment capability of LLMs and applied age gaps to identify proteomic biomarkers associated with accelerated aging and to develop risk prediction models of 270 diseases. Interpretability analyses were also conducted to explore the decision-making process of LLMs. In conclusion, our LLM-based aging assessment framework offers a precise, reliable and cost-effective approach for estimating overall and organ-specific aging. It has potential for personalized aging assessment and health management in large-scale general populations.

Aging is a major risk factor for mortality and chronic diseases, imposing a significant health burden on society^{1–3}. Previous studies have indicated that aging is an intricate, multidimensional process^{4,5}, exhibiting significant heterogeneity at the body^{6,7}, organ^{8–10}, molecular^{11,12} and other levels¹³, influenced by various factors such as environment and genetics¹⁴. For clinical practice, an overall aging proxy provides a better assessment of comprehensive health, whereas organ-specific aging proxies offer insights into health of specific organs. All proxies cannot

be adequately represented by chronological age. Hence, it is imperative to develop overall and organ-specific aging proxies to more precisely reflect the progression of multidimensional aging to inform health risks and potential interventions^{15,16}. Despite this need, the development of practical methods to accurately and conveniently assess aging for large-scale general population remains a critical challenge¹⁷.

Previous studies have developed various aging proxies, such as the frailty index and biological ages. The frailty index assesses overall

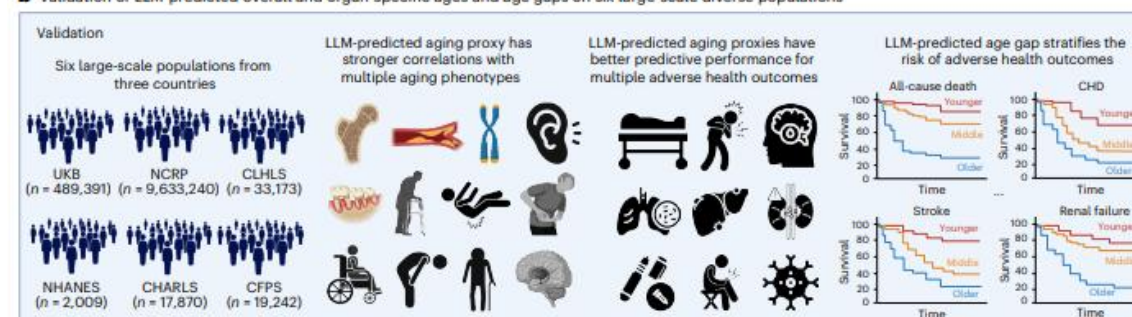
✉ e-mail: mawz@tsinghua.edu.cn; yangyn5126@163.com; qlandi@tsinghua.edu.cn

Nature Medicine

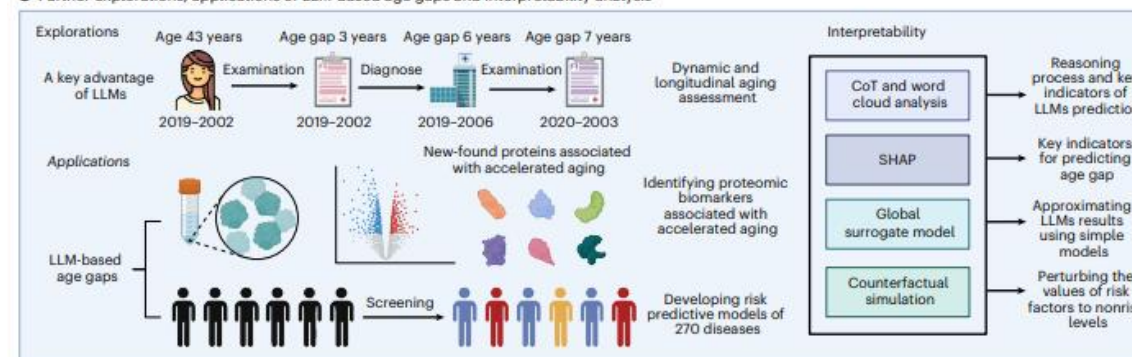
a The framework of predicting the overall and organ-specific ages using large language models



b Validation of LLM-predicted overall and organ-specific ages and age gaps on six large-scale diverse populations



c Further explorations, applications of LLM-based age gaps and interpretability analysis





医薬品と医療機器開発のベースの違い

医薬品→ベースはシーズドリブン



医療機器→ベースはニーズドリブン



医療機器開発は徹底した現場ニーズ把握が重要



現場ニーズ

徹底した現場のヒアリング・Must haveになるか？

ソリューション

テクノロジーベースで解決可能かどうかを検証

実現可能性とマーケットサイズ

製造可能かを検証すると共にマーケットを意識

競争優位性

差別化できるか、保険収載などで独占できるか

AIできた！ではなく 徹底した現場ニーズが重要

現場ニーズ

Must haveになって現場のフローを変えられるか？

ソリューション

AI開発のデータや技術が入手可能か 計算性能は十分か

実現可能性とマーケットサイズ

持続可能なシステムで提供可能か、AIのマーケットは新規か置換か

競争優位性

他社が真似しにくいAIか、AIとしての保険収載/加算を得られるか

- 技術面：開発時の精度と臨床要求の乖離
- 規制面：承認・再学習ルールの未整備
- 経済面：導入コスト・保険償還制度の課題
- 社会面：説明責任・臨床データ共有の難しさ

一方、臨床要件として重要なのは以下の点を明確にすること

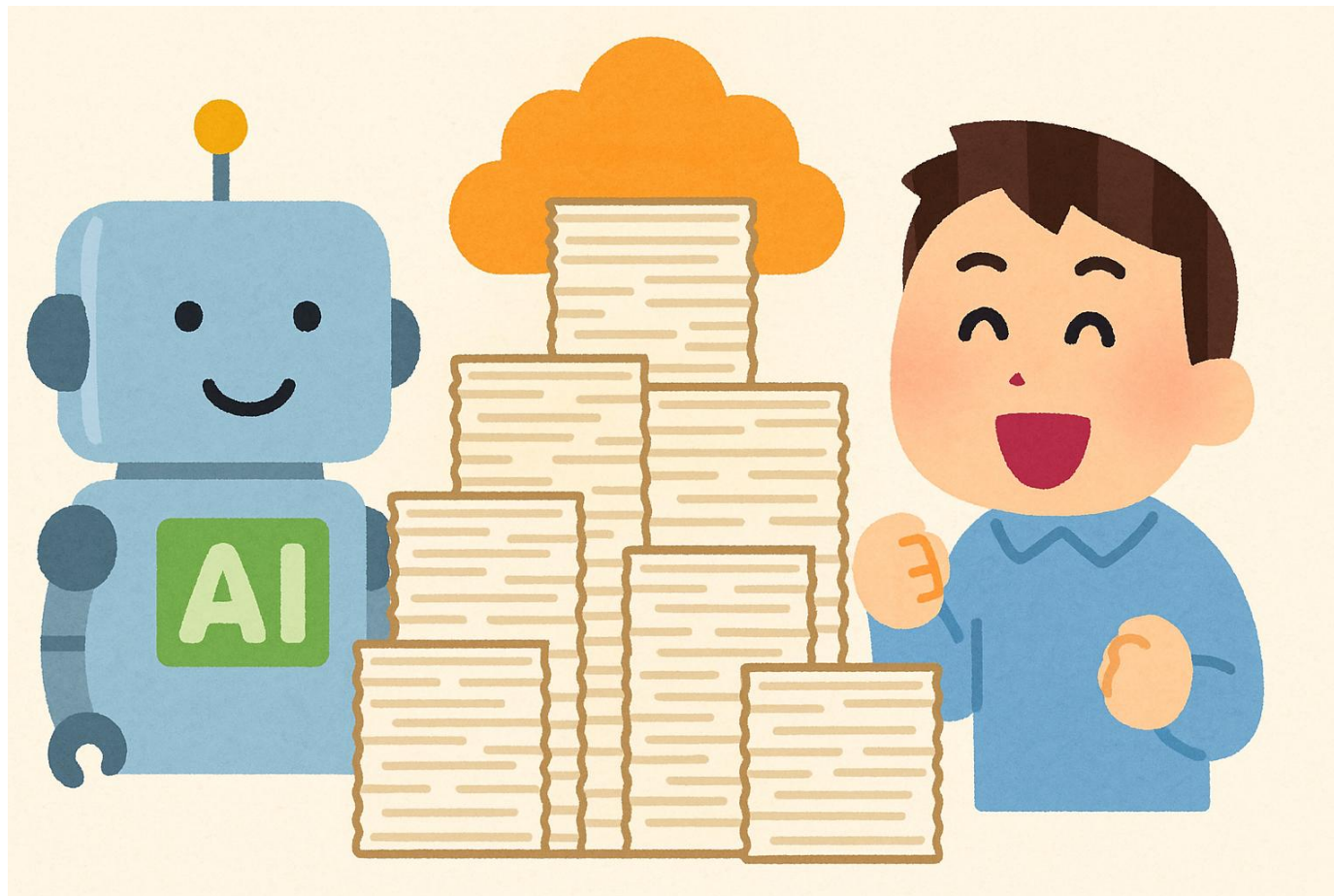
- どのような診療場面で使用するか
- 検査前確率を考慮した上で必要精度はどの程度か
- 使用後の臨床アクションが明確か

感度98% 特異度95%の検査があるとしています

検査前確率	検査が陽性の際の的中率 (PPV)	検査が陰性の際の的中率 (NPV)
20%	83.1%	99.5%
0.1%		99.9%

AIを用いたSaMDはデータ量が命？

- ほとんど手に入らないデータ(ex. 稀少疾患のプロテオーム)は希少
- それ以外のものは...



1000万件超の巨大心電図はすでにモデル化



ORIGINAL ARTICLE



An Electrocardiogram Foundation Model Built on over 10 Million Recordings

Authors: Jun Li, B.S. , Aaron D. Aguirre, M.D., Ph.D. , Valderly Moura Junior, Ph.D. , Jiarui Jin, B.S. , Che Liu, M.S. , Lanhai Zhong, B.S. , Chenxi Sun, Ph.D. , Gari Clifford, Ph.D. , M. Brandon Westover, M.D., Ph.D. , and Shenda Hong, Ph.D.  [Author Info & Affiliations](#)

Published June 26, 2025 | NEJM AI 2025;2(7) | DOI: 10.1056/Aloa2401033  | VOL. 2 NO. 7 | Copyright © 2025

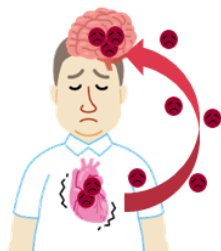
- 圧倒的なスケールと多様性を誇るデータセット
 - 約180万人の被験者から>1000万件の12誘導ECG
 - 150種類にも及ぶ診断ラベルを教師付きで学習
 - 微調整(fine-tuning)して多様な下流タスクに応用できる

時代は ”どうデータを抑えるか” から
”どの現場にどのようなソリューションを展開するか”
という実装ベースの現場展開力に移行

なぜ心房細動スクリーニングをAIの対象に？

根幹にある社会問題

致死的な脳卒中を起こす 心房細動の見逃し



- 心房細動は未治療で7~8人に1人が死に至る重篤な脳梗塞を起こす
- 推定患者国内200万人の約半数が未診断・未治療で放置されている
- 心房細動による脳梗塞は年間数万人に達すると見られる

医療機関が抱える問題

膨大な波形データの DX化の遅れ



- クリニックの半数ほどが紙媒体で診療データを管理
- 諸外国と比較してDXが遅れている

高コストによる 診断機会の損失



- 検査技師・医師の作業負担・コストが高い一方で保険償還価格は同一なため心房細動の発見に有効な長時間心電図検査の導入が進んでいない

未発見の心房細動をスクリーニングして脳梗塞をAIで激減させる



AIクラウド型SaMDを用いて医療現場の作業効率改善を通して
脳梗塞の主原因である心房細動のスクリーニングのマネタイズと普及を目指す

心房細動は高血圧・糖尿病と並ぶ危険因子

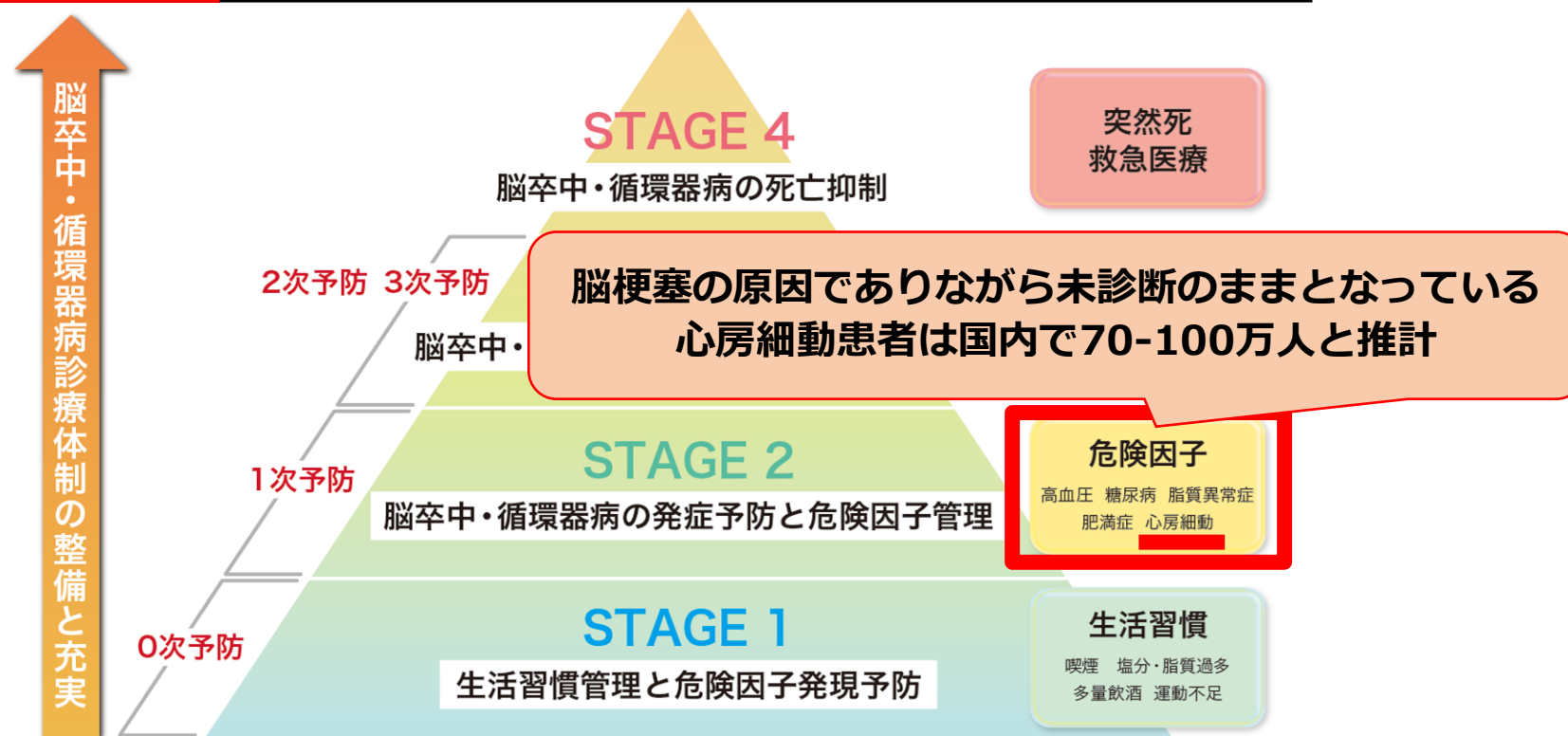
脳卒中と循環器病克服5力年計画

～日本脳卒中学会/日本循環器学会～ 第1次(2015-), 第2次(2020-)



戦略4：予防・国民への啓発

心房細動は高血圧・糖尿病・脂質異常症・肥満と並ぶ危険因子



脳卒中と循環器病克服5力年計画資料(http://www.j-circ.or.jp/five_year/)より

世界初: “これから起きる心房細動発作”を予測するAI



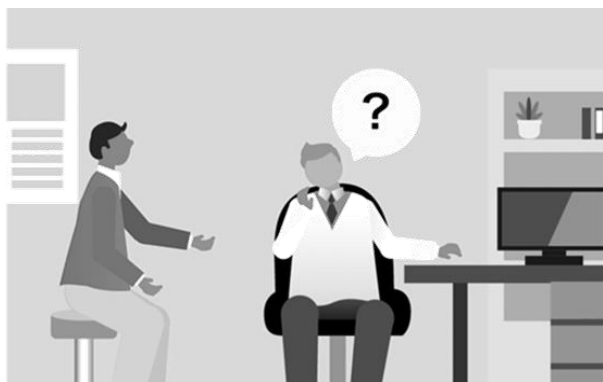
新しいAI医療機器『Smart PAFin』が薬事承認取得

発作を起こす可能性 “**隠れ心房細動**” が検出可能に
➡ 心房細動の早期発見が脳梗塞の撲滅へ

従来の不整脈の診断技術



発作がある時
にしか診断できない



Smart PAFin



発作がない時
でも発作の兆候検出が可能





ディープラーニング②

PAF発作兆候判定機

長時間心電図解析ソフトウェア SmartRobin AI シリーズ



心電図解析をAIとともに

～ヒトとAIのタスクシェアで医療現場の負担を削減～

- ・ 従来の解析装置をディープラーニング技術を用いたAIで置換し、認証で社会実装
- ・ 長時間心電図検査をAIで効率化する“現場”を創る役割



SaMDをクラウド化しAIと融合

従来の医療機器



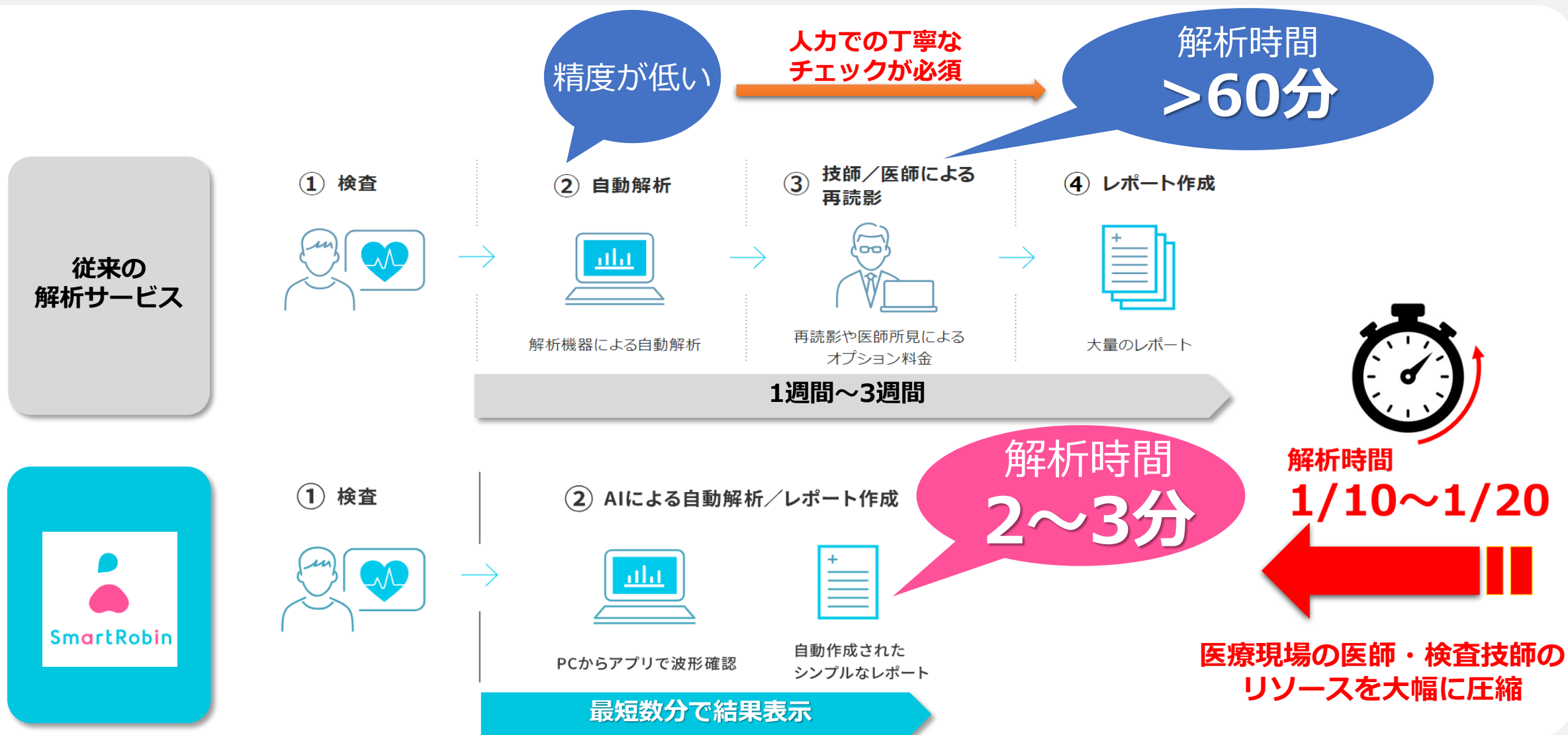
- 高価な減価償却費
- その場所でのみしか使えない
- 使わないリスクも高い
- アップデートが大変
- 人が確認することを前提に設計

クラウド提供SaaS型ソフトウェア医療機器



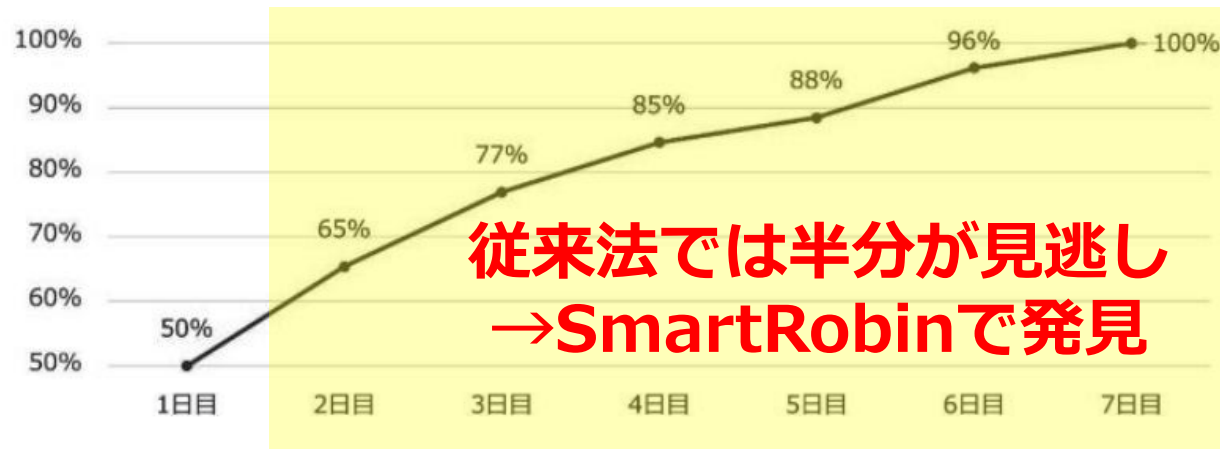
- 初期導入コストが安い
- どこからでもアクセス可能
- 使用実績に基づき課金
- 最新のアプリケーション提供が可能
- AI実装も容易

AI-SaMDにより医療現場DXを推進し市場を作る



SmartRobinで現場DX→心房細動発見が倍増

従来の24時間ソリューションと比較して
SmartRobinで心房細動発見が倍増
→脳梗塞予防につながるエビデンス創出



臨床現場で使用した実績：
SmartRobinのAF検出感度は100%

n = 100	感度 (%)	特異度 (%)
全集団	100	68.9



ステートメントに唯一収載されたAI医療機器



>150の医療機関に導入

2024年1年間で+300%成長
収集データも同意下で利活用可能



累計オーダー数

10,000件突破!



エキスパートコンセンサスステートメント

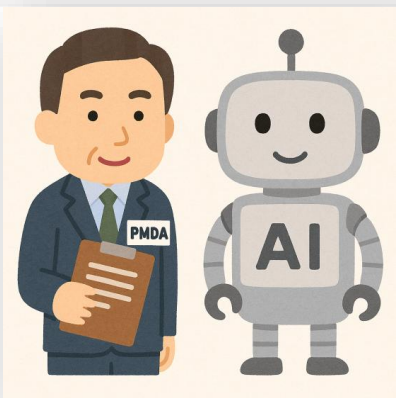
2024年日本不整脈心電学会 / 日本循環器学会
携帯型 / 装着型心電計の適切使用に関する
コンセンサスステートメント

〈執筆委員〉池田隆徳¹(班長) 芦原貴司² 岩崎雄樹³ 大野真紀⁴ 鍵山暢之⁵ 木村雄弘⁶
草野研吾⁷ 河野律子⁸ 朔 啓太⁹ 笹野哲郎¹⁰ 妹尾恵太郎¹¹ 高月誠司¹² 高橋尚彦¹³
高見 充¹⁴ 中野由紀子¹⁵ 橋本賢一¹⁶ 藤生克仁¹⁷ 藤野紀之¹⁸ 水野 篤¹⁹ 吉岡公一郎²⁰
渡邊英一²¹

〈外部評価委員〉清水 渉²² 野出孝一²³

- 心電図データの解析は各社独自のソフトウェアによって行われるが、**複数機種に対応し、長時間心電図記録から心房細動診断に特化したAI クラウドベースのプログラム（長時間心電図解析ソフトウェアSmartRobin AI シリーズ）も使用されている。**

循環器の医療現場でSaaS型のAI医療機器のマーケットを創造



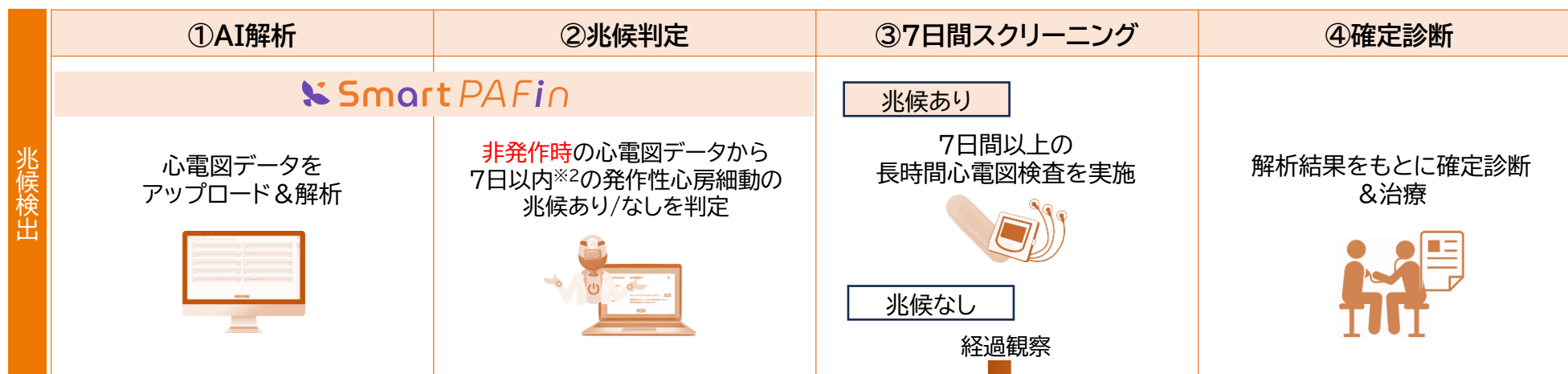
- どのような対象に使いますか？
- どの程度の精度が出れば臨床的に意味がありますか？
- 判定後のフローはどうなりますか？

治験では評価後7日間以内に実際に心房細動がどうかを検証
検査陽性者・陰性者での性能をそれぞれ評価した

大切なことは

- AIが一定の精度を出せば承認されるというものではない
- 臨床上のユースケースを明確化し、臨床試験で求められる精度達成のエビデンスを取りに行く

治験で検査後7日間以内に実際に心房細動がでた
症例での検出力を検討し感度91%*で検出



高感度なので陰性的中率が高い

発作を起こすタイミングまで評価
治験時から AI判定後の臨床フローを意識することで
その後の“隠れ心房細動”検出の臨床的アクションが明確化が可能に

心房細動確定診断のため2製品でクロスセル効果

「PAFinで予測して、Robinで捉える！」

SmartPAFin

- ✓ **2** 誘導以上のホルター心電計で
- ✓ **5** 分間の心電図データを解析し
- ✓ **7** 日以内の心房細動の発作を予測



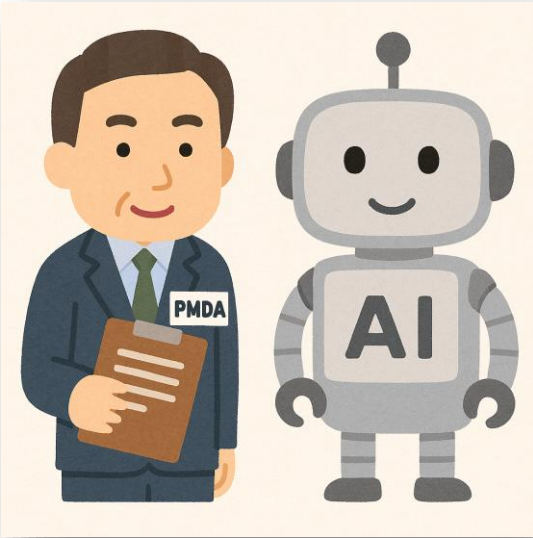
※画面はイメージです

SmartRobin

- ✓ **95**%心房細動の検出精度で
- ✓ **24**時間分の長時間心電図データを
- ✓ **3**分で解析完了



まとめ：AI-SaMDを社会実装するには



- AI-SaMDはシーズドリブンであってはいけない、ニーズドリブンで現場にマッチするかどうかを実装面でのポイント
- “AIが使われる＝その医療現場のフローを変える事”なのでそのインセンティブが現場にあることが大切
- AI-SaMDの承認のハードルは規制側が決めるものではなく開発者側が決めるもの
- 具体的なユースケースを徹底的に調査し、そこにハマる設計と精度を実現することが承認の王道