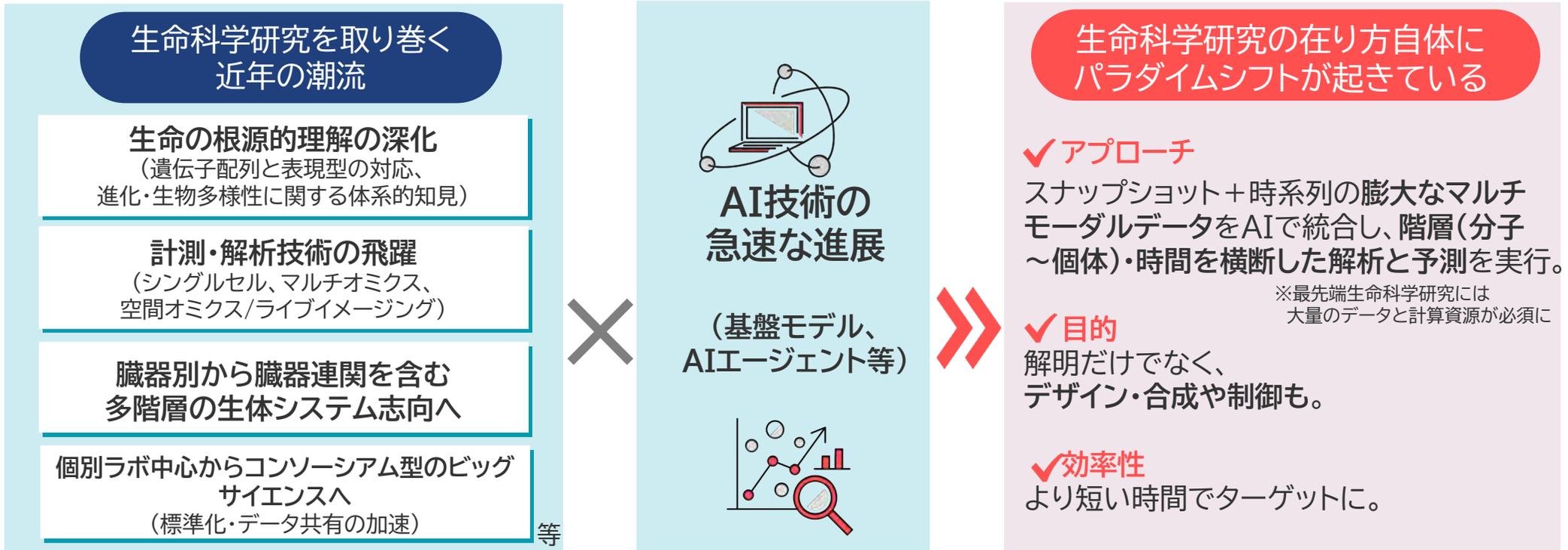


医薬品開発におけるAIの活用について 関する文部科学省の取組について

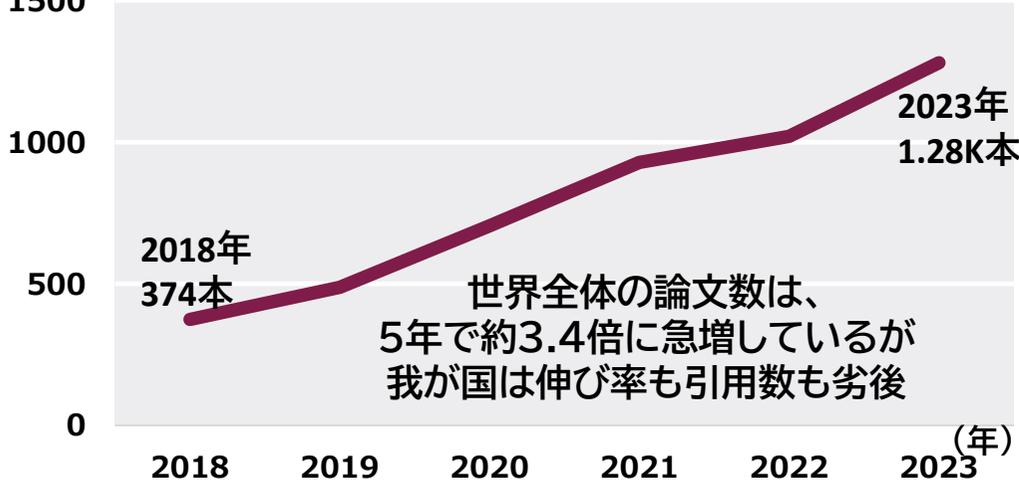
令和8年2月27日

第2回医薬品開発協議会

医薬品開発におけるAIの活用に関する状況と課題



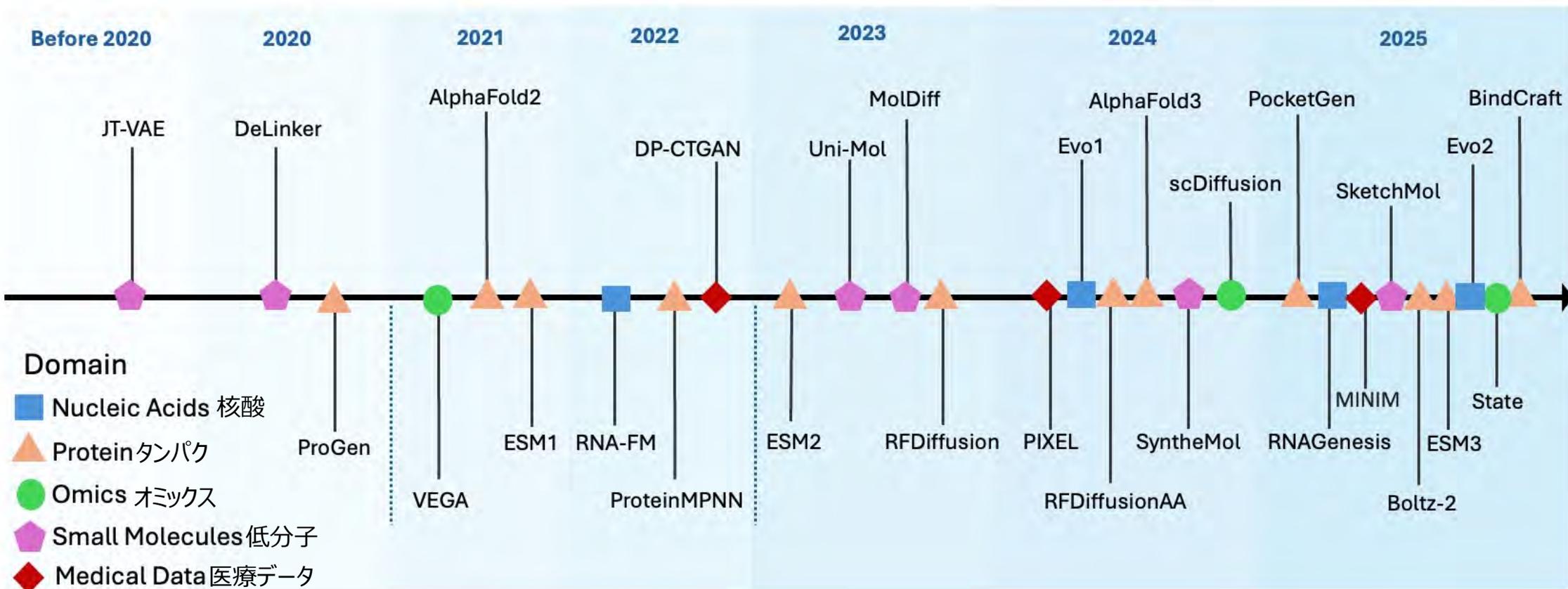
(論文数) 【AI×Pharmacology の世界全体の論文数推移】



国	論文数 (2018-2023)	伸び率 (2018-2023)	世界シェア	1論文当たり平均引用数
中国	1,326	302%	28%	27.9
米国	1,050	145%	22%	40.3
インド	517	573%	11%	18.8
英国	220	213%	5%	61.4
韓国	170	917%	4%	32.1
ドイツ	148	192%	3%	56.0
カナダ	142	110%	3%	39.0
日本	140	56%	3%	20.7
フランス	116	63%	2%	22.6
豪州	95	100%	2%	21.3

(参考)様々なライフサイエンス分野のAIモデル開発の状況

- ◆ 2021年のAlphaFold2の登場が大きな転換点となり、ライフサイエンス分野におけるAI(特に深層学習や基盤モデル(Foundation Model))を活用した研究が急増。
- ◆ さらに2023年の生成AIブームを受けて、大規模言語モデル(LLM)の医療応用、拡散モデル(Diffusion Model)のバイオ分野での応用、ラボオートメーションの研究なども加速。



(出典) Zaixi Zhang et al. (2025, Oct) *Generative AI for Biosciences: Emerging Threats and Roadmap to Biosecurity*
<https://arxiv.org/abs/2510.15975>

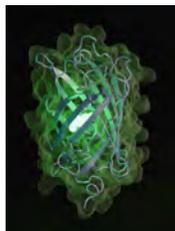
(参考)欧米のライフサイエンス分野におけるAI活用に関する研究開発動向

生命・医学におけるAI基盤モデルの開発

タンパク言語モデル

元Meta社の研究者らが設立したEvolutionaryScale社が約28億のタンパク質を学習した**タンパク言語モデル(ESM3)**を開発。タンパク質の配列、構造、機能の3つを同時に推論し自然界にはない新しいタンパクを生成することも可能に。

【出典】2024年6月25日, EvolutionaryScale社,
「Simulating 500 million years of evolution with a language model」



新しい緑色
蛍光タンパク質

タンパク・デザインモデル

スイス・ローザンヌ工科大学を中心に、MITやオランダの研究チームが連携する国際チームが、AlphaFold2の技術を応用し、標的タンパク質の特定領域に結合し、その機能を制御するタンパク質を、従来よりも高成功率で設計できるモデル(**BindCraft**)を開発。

【出典】2025年8月27日, Martin Pacesa et al., Nature 「One-shot design of functional protein binders with BindCraft」

ゲノム言語モデル

米・Arc Institute(非営利研究機関)が、スタンフォード大、UCバークレー、UCサンフランシスコ、NVIDIAと連携し、微生物、植物、動物、ヒトから集めた約9.3兆のDNA塩基対の情報を学習した**ゲノム言語モデル(Evo2)**を開発。遺伝子変異がタンパク質やRNAに及ぼす影響の評価や、新しいDNA配列の生成も可能に。

【出典】2025年2月19日 Arc Institute 「AI can now model and design the genetic code for all domains of life with Evo 2」

予測モデル

Google DeepMind社が、非コード領域を含む最大100万塩基対のDNA配列を解析し、遺伝子発現、スプライシングパターン、クロマチンの特徴など多様な分子情報(モダリティ)を単一塩基対レベルで同時に予測可能な、ヒトやマウスのゲノムで学習したモデル**AlphaGenomics**を開発。遺伝子変異の影響を、効率的に短時間で予測することを可能に。

【出典】2025年6月25日, Google DeepMind, 「AlphaGenome: AI for better understanding the genome」

ゲノム言語モデルとLLMの統合モデル

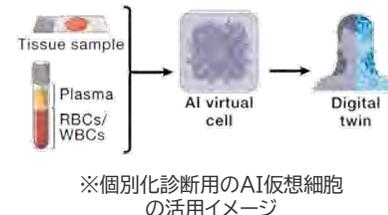
トロント大学等の研究チームが、ゲノム言語モデルと大規模言語モデル(LLM)を統合し、理解可能なモデル**BioReason**を開発。

【出典】2025年5月29日, Adibvafa Fallahpour et al, arXiv:2505.23579
「BioReason: Incentivizing Multimodal Biological Reasoning within a DNA-LLM Model」

AI仮想細胞モデルの開発

✓米・Chan Zuckerberg Initiativeは、2023年9月、**AI仮想細胞(AI Virtual Cell)計画**を明らかにし、2024年12月にはスタンフォード大学等と、**AI仮想細胞基盤モデル**の具体構想を提案

【出典】2024年12月12日, Charlotte Bunne et al, Cell,
「How to build the virtual cell with artificial intelligence: Priorities and opportunities」



✓米・イェール大、Google DeepMind等が、シングルセルデータ(scRNA-seq)を人間が理解できる言語として扱う**仮想細胞モデルC2S-Scale**を開発。「この細胞は薬Xにどう反応するか?」といった質問に、生物学的情報に基づいた回答が自然言語で得られる。

【出典】2025年4月17日, Syed Asad Rizvi et al,
「Scaling Large Language Models for Next-Generation Single-Cell Analysis」

✓米・Arc InstituteとUCバークレー、スタンフォード大、UCサンフランシスコ、ペンシルベニア大、イェール大が、幹細胞や、ガン細胞、免疫細胞が薬物、サイトカイン、遺伝的摂動にどう反応するかを予測する**仮想細胞モデル(STATE)**を、約1.7億個の細胞からの観察データと、1億個以上の細胞からの摂動データを学習させて開発。

【出典】2025年6月23日, Arc Institute
「Predicting cellular responses to perturbation across diverse contexts with State」

戦略的な取組を進める研究機関

✓米・ブロード研究所の“**Eric and Wendy Schmidt Center**”開設

2021年3月、機械学習と生物学を融合し、生命のプログラムの理解を目指す、センターを1.5億ドルの寄附を得て開設。タンパク質の細胞内局在予測モデル等の開発が進められている。

✓米・ハーワード・ヒューズ医学研究所(HHMI)の『**AI@HHMI**』計画

2024年8月、科学的発見の加速のため、AIを活用した生物医学研究に10年間で5億ドル規模を投資することを発表。ジャネリア・リサーチ・キャンパスを中心に、HHMI全体でAIの活用を促進。

✓欧州分子生物学研究所(EMBL)の『**Science AI Strategy**』

2025年2月、欧州の生命科学分野のAI活用の変革を目指した戦略を提示。生物学的ドメインに特化した基盤モデルの開発や、欧州の大規模バイオデータのAI対応用に整備すること等が掲げられている。

医薬品開発におけるAIの活用に関する文部科学省の取組



文部科学省

- 文部科学省においては、全国の研究者に対して**先端研究基盤やAI創薬支援を提供**することで研究成果の実用化を促進するとともに、大学・研究機関等において**科学研究におけるAIの利活用（AI for Science）を推進**。
- 特に「最先端のデータを創出する実験科学」、「良質なデータを取得する計測技術」、「高品質なデータアセット」等の我が国の強みを活かしながら、高精度・高効率な医薬品開発にも資する、大規模なデータ取得を通じたバイオ生成基盤モデルの開発等を推進。

AI for Scienceの推進

- **科学研究におけるAIの利活用（AI for Science）**により科学研究の在り方そのものが変革されつつあり、**米国・EU等は国家的な取組として戦略的に推進**（ターゲットにはバイオテクノロジー、創薬も含まれる）する中、我が国においても年度内に**AI for Scienceの推進に向けた基本的な戦略方針を策定**するための検討を実施。
- **戦略方針を踏まえ、AI for Scienceによる科学研究革新プログラム（令和7年度補正予算370億円）等の新たな支援を開始予定**。

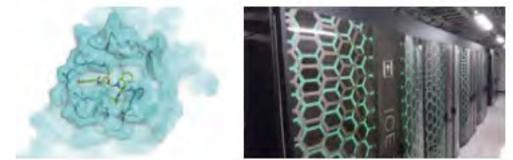


米国「ジェネシス・ミッション」と日米連携（文部科学省とDOEは「協力のための意向表明（Statement of Intent : SOI）」に署名）

AMED 生命科学・創薬研究支援基盤事業（BINDS）



- 創薬・ライフサイエンス分野における先端的な機器整備の実現を通じて研究支援技術の高度化を図り、生命科学・創薬研究成果の実用化を促進する事業。
- BINDS内の**インシリコユニット**において、**候補化合物を探索するインシリコスクリーニング**や**有望化合物の活性・薬物動態・安全性プロファイルの最適化設計**などのAI創薬支援を実施。



インシリコ解析ユニット

理化学研究所 創薬・医療技術基盤プログラム（DMP）



- **理化学研究所でのライフサイエンス研究で培われた研究基盤を活用し、日本発の革新的な医薬や医療技術の創出**を目指すプログラム。
- **基礎研究に基づく高品質なデータと理研の計算科学基盤（スーパーコンピュータ等）を活かし、AI 予測モデルや分子生成 AIなど多様な創薬 AI 技術を総合的に開発**。



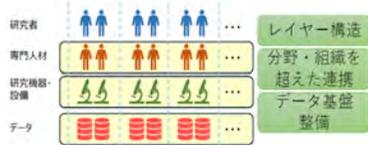
「富岳」量子コンピューターの活用

AI for Science の推進に向けた基本的な戦略方針（概要イメージ）

資料3
AI for Science推進委員会
(第1回)
令和8年2月9日

第7期科学技術・イノベーション基本計画の方向性

- ▶ 科学技術推進システムを刷新し、科学技術政策を大転換
 - 現状認識の一つとして… **AIと科学の融合による研究開発パラダイムの転換**
- 2章 知の基盤としての「科学の再興」
- 4.AI for Scienceによる科学研究の革新



AI基本計画（R7.12閣議決定）

- ▶ イノベーション促進とリスク対応の両立
- ▶ **信頼できるAI**の追求
- ▶ **世界で最もAIを開発・活用しやすい国**を目指す

海外動向

文部科学省

- 各国は**AI for Science**を**国家的ミッション**と位置づけ
- **研究投資、計算基盤整備、人材育成等を強化**
- 科学とビジネスが近接化し「**勝者総取り**」構造
- **科学的発見のプロセス**自体が、学際的、データ駆動型、計算集約型へと**大きく変化**

日本の強み

- ▶ **情報基盤**：世界最高水準の情報流通基盤（SINET）・研究データ基盤（NII RDC）・計算基盤（富岳・富岳NEXT・HPCI等）
- ▶ **研究基盤**：世界トップレベルの基礎科学力と多様な研究者層、世界最先端の研究装置群と大型研究施設、信頼性の高い実験・観測データの蓄積
- ▶ **社会基盤**：世界有数の経済規模、精密な製造・計測技術・ロボティクス、すり合わせや暗黙知を含む現場知、AIに対する社会的・産業的な需要

目的 I. 科学研究の革新と科学的発見の加速・質の変革、II. 研究力の抜本的強化と科学の再興、III. 国際的優位性・戦略的自律性の確保

中長期的な取組目標 **科学基盤モデル/エージェントやAI駆動ラボの活用により重要技術領域の先端的成果創出及び研究開発期間を1/10に**

今後5年間の目標 **AI for Scienceの推進により、日本の科学研究における国際優位性の確保**

（ターゲット例）

- **3年後までに、新素材開発速度10倍の潜在力を有するAI駆動ラボシステムを開発。**
将来は、AI駆動ラボシステムを用いて、我が国企業が国際的サプライチェーン上不可欠なマテリアルを量産する。
- **3年後までに、大規模なデータ取得を通じて、高機能なバイオ製品の高効率設計を実現するバイオ生成基盤モデルを開発。**
将来は、仮想細胞・生体モデルや、植物、動物、ヒト・臓器等の“デジタルツインモデル”を実現し、高精度かつ高効率なバイオ製品開発や創薬等に貢献する。
- **3年後までに、AIEージェント群による、最先端大型研究施設・研究装置からの大量・高品質データ産出や、仮説検証・実験の自動化・自律化を実現。**
新規性の高い研究を探索的に行うシステムの開発を通じて、科学研究の新しい方法論を示す。

戦略的な国際連携
(米国・英国など)

世界を先導する
科学研究成果の創出

AI for Science の波及・振興
による科学研究力の底上げ

AI for Science を支える
研究基盤の構築

（具体的な取組内容）

① **研究力・人材**
AI研究者等の育成
×
AI活用促進

② **計算資源**
戦略的増強
×
利便性向上

③ **研究データ**
高品質データの創出
×
データの一元化

- AI for Science のあらゆる分野での波及・振興と日本の強みを生かした重点領域の設定・投資を両輪で推進、世界トップ層との戦略的国際共同研究を推進
- AIの基礎研究含むAIそのものの研究の強化（リスク対応含む）
- 国際連携・産学連携を通じ、AI・計算資源・データに精通した人材の参画・育成、技術専門職の育成・確保、評価や処遇の見直し
- 世界最高水準の次世代AI・HPC融合プラットフォーム「富岳NEXT」の開発
- AI共用計算資源の戦略的な増強と利便性（機動性、アクセス性、相互運用性）の向上
- 産業界との連携及び国際連携を通じた計算資源の有効活用
- 戦略的価値の高いデータセットの特定・構築
- 自動化・自律化した研究設備等の整備と研究データ創出プロセスの標準化
- AI時代に即した次世代情報基盤の構築・活用、データの一元的運用

参 考



課題・取組の方向性

- ▶ タンパク質の構造予測を行うAlphaFold（ノーベル賞）は研究にかかる時間とコストを劇的に削減するなど、**AIは、研究力の生産性の向上のみならず、科学研究の在り方そのものを変革**。国際的にAIの研究開発や利活用への投資が進む中、**自国でAI研究開発力を保持することは安全保障上極めて重要**。科学研究におけるAI利活用（AI for Science）において、米国・EU等は国家的な取組として、リソース（計算資源・研究資源・人材・データ等）を有効活用し、戦略的に推進。
- ▶ 我が国においては、世界最高水準の情報基盤を有するとともに、**ライフ・マテリアル等の重点分野において次のAI開発・利活用の要となる質の高い実験データを持つ等の強み**を有しており、これらのリソースを最大限活用し、**科学基盤モデル・AIEージェント開発、次世代AI駆動ラボシステム開発、これらの実装に向けた取組を進めることで、第7期科学技術・イノベーション基本計画で目指す研究力向上を牽引**。

事業内容

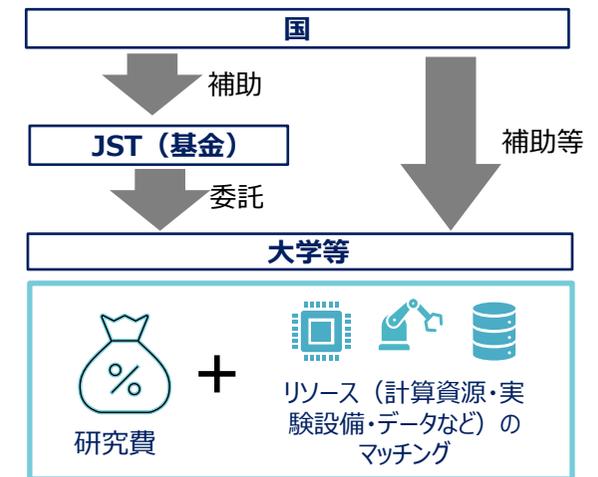
事業実施期間 ~令和10年度

- 国のコミットメントの下で、我が国が有する**計算資源等のリソースを戦略的かつ機動的に配分しながら**、重点領域への集中投資により世界をリードすることを目指す**プロジェクト型（基金事業）**と、あらゆる分野における波及・振興及び先駆的な研究を目指す**チャレンジ型**を**両輪**とし、**AI for Science先進国**の地位を確立する。

- ① **プロジェクト型**：我が国の勝ち筋となる**重点領域**において、シミュレーションデータに加え、実験データの取得・活用による我が国発の最先端AI基盤モデル・AIEージェント開発、次世代AI駆動ラボシステム開発、これらの実装に向けた取組を一体的に推進。我が国の研究力を抜本的に強化するとともに、産学の協働により、研究開発投資を促進し、先駆的取組の早期実装・ビジネス化により**科学研究を変革するイノベーション**を創出。
- ② **チャレンジ型**：あらゆる分野の研究者がAIを活用して科学研究の高度化・加速化を図るため、計算資源の確保等の研究環境を整備し、**アカデミア全体にAI for Scienceの波及・振興を促進し、意欲ある研究者による次の種や芽となる新たなアイデアへの挑戦への支援**を行うとともに、我が国独自の競争優位を築く先駆的な研究を創出。

※上記の他、AI for Scienceに不可欠な計算基盤の環境整備として、76億円を別途計上。

【事業スキーム】



【取組のイメージ】

AI×実験科学 = ライフサイエンスの再興
 <アセット>
 ・最先端データを創出する実験科学
 ・良質なデータを測る技術
 ・データアセット・バイオリソース

×AI

・バーチャル臨床試験
 ・個別化診断
 ・創薬・医療

創薬・精密医療・バイオものづくり等の新産業創出

AI×装置×産学知 = マテリアル開発の革新
 <アセット>
 ・ラボから量産まで一気通貫の開発・実装能力
 ・世界有数の実験データベース&産業界の暗黙知データ
 ・先端的な計測技術と国内機器産業クラスター

×AI

・オンデマンド材料設計
 ・自律ラボで未知材料を自動探索

国内外から投資が集まり、短期間で革新的マテリアルが量産可能となるR&D拠点群を形成

AI×多様な分野 = 新たな日本の勝ち筋の探究
 ・AI for Scienceの波及・振興を促進するとともに、あらゆる分野の意欲ある研究者による新たな勝ち筋の創出

×AI

量子 数理物理学 認知科学 都市工学 農業 考古学 フェージョン エネルギー等

「プロジェクト型」	「チャレンジ型」
320億円	50億円
・支援件数：5領域×3チーム ・支援規模：20億円程度/件 ・支援期間：原則3年	・支援件数：1,000件程度 ・支援規模：500万円程度/件 ・支援期間：～1年

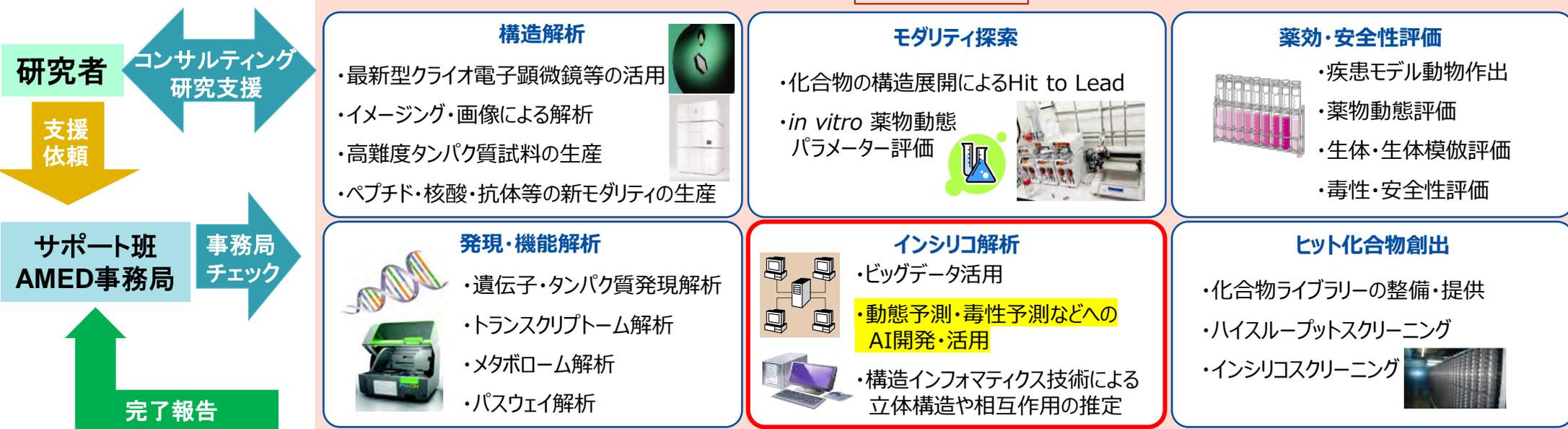
現状・課題

- 「健康・医療戦略」(令和7年2月閣議決定)に基づき、広くライフサイエンス分野の研究発展に資する高度な技術や施設等の先端研究基盤を整備・維持・共用することにより、大学・研究機関等による基礎的研究成果の実用化を促進。
- また、本事業は「創薬力の向上により国民に最新の医薬品を迅速に届けるための構想会議中間とりまとめ」(令和6年5月22日)において、非臨床試験や共用・基盤整備の推進に活用されることが記載されている。
- 令和8年度においては、大規模解析の効率化・高速化のための機器整備などを行うことにより、「経済財政運営と改革の基本方針2025」(令和7年6月13日閣議決定)や「新しい資本主義のグランドデザイン及び実行計画2025年改訂版」(令和7年6月13日閣議決定)で示されている、先端研究設備・機器の戦略的な整備・共用・高度化の推進や創薬力の抜本的強化のための研究支援基盤整備を図る。

事業内容

- 創薬・ライフサイエンス分野における先端的な機器整備の実現を通じて研究支援技術の高度化を図り、生命科学・創薬研究成果の実用化を促進する。
- 関係機関が連携し、高度な解析機器を効果的かつ効率的に運用できる人材の育成を推進する。

支援ユニット



※上記取り組みに加え、各ユニットの機器を利用した人材育成の推進を行う

事業実施期間 令和4年度～令和8年度

交付先 国公立大学、研究開発法人等の国内研究機関

【事業スキーム】



(担当：研究振興局ライフサイエンス課)

- 世界中でAIの研究開発や利活用への投資が進んでおり、各国はAIを戦略的に重要技術と位置づけ、**AIに関するインフラ整備・研究投資などを総合的に進める国家戦略**を整備している。
- 最近では、米国やEU等において**AI for Scienceに関する取組**が強力に進められている。

米国

「America's AI Action Plan」(2025.7)

①AIイノベーションの加速、②AIインフラの整備、③国際的な外交・安全保障での主導の3本柱で構成する包括的国家戦略。

<AI for Scienceに関する主な取組>

- ✓ 科学、安全保障、技術のためのAIフロンティア (FASST)
- ✓ AI研究のためのインフラ提供 NAIRR Pilot

「GENESIS MISSION」(2025.11)

EU

「AI大陸行動計画」(2025.4)

EUが「AI大陸」としてAI分野の世界のリーダーとなることを目指し、インフラ、データ、人材、応用、規制の5分野で包括的に推進する計画を示した。

「欧州におけるAI in Science戦略」(2025.10)

仮想的な研究機関「Resource for AI Science in Europe (RAISE)」を構築し、計算資源、データ、ノウハウ、人材、研究資金などのAI資源を一元化させ、研究の効率と質を高める。

<AI for Scienceに関する主な取組>

- ✓ 計算資源とデータ・人材の集積拠点AIファクトリー/AIギガファクトリー
- ✓ 欧州データ統合戦略(策定予定)

英国

「AI機会行動計画:政府回答」(2025.1)

基盤整備・生活変革・国産AI保護の3領域を柱に、研究資源強化や特区設置、データ整備、人材育成、公共部門導入、官民連携を推進する方針を示した計画。

「英国AI for Science戦略」(2025.11)

英国が強みを持つ5つの分野をターゲットとして、AI駆動科学の加速・AIによる科学研究の変化に関する研究への投資、データ・計算資源・人材と研究文化に関する取組を実施する。

<AI for Scienceに関する主な取組>

- ✓ 学術向けAI計算基盤 AIRR
- ✓ 創薬データ基盤OpenBindコンソーシアム

中国

「新世代人工知能開発計画」(2017.7)

2030年までの三段階目標を掲げ、理論と融合研究を推進する国家AI戦略。

「『人工知能プラス』行動のさらなる実施に関する意見」(2025.8)

2035年までの三段階目標を掲げ、AIを社会・経済全域に深く融合し新質生産力と知能社会を育成する行動提言。

<AI for Scienceに関する主な取組>

- ✓ AIを活用して科学研究や技術開発を加速・高度化する「AI+科学技術」

■米・GENESIS MISSION

- ✓ AIによる科学研究と技術革新の抜本的改革を目指す国家プロジェクト
- ✓ **10年間で米国の科学研究および技術革新の生産性と影響力を2倍にする**
- ✓ **中核的要素**: American Science and Security Platformの構築、政府保有科学データのAI利活用、産官学の協働体制
- ✓ **主要課題領域**(エネルギー覇権、科学的発見の加速、国家安全保障の確保)
- ✓ 2025年12月DOEが**3.2億ドル超**の初期投資を発表

■英・AI for Science Strategy

- ✓ 科学的発見プロセスそのものを革新
- ✓ **3つの柱**(データ、計算基盤、人材・文化)
- ✓ **5つの重点分野**(先端材料、核融合、医療研究、エンジニアリング・バイオロジー、量子技術)
- ✓ **15の具体的アクション**(AI駆動科学促進、データのFAIR原則の義務化、信頼性や環境負荷低減など)
- ✓ **最初のミッション**: **2030年までにAIを活用して「試験開始可能な薬物候補を100日以内に創出」**
- ✓ 2026~2030年に**約1.37億ポンド**を投資

- 11月24日、**トランプ大統領**は、発見科学を加速させ、国家安全保障を強化し、エネルギーイノベーションを促進するために、**世界で最も強力な科学プラットフォームを構築するための国家的なイニシアチブ「GENESIS MISSION」の開始を指示する大統領令**に署名。
- エネルギー省(DOE)に、スーパーコンピュータと独自のデータ資産を統合し、科学的基盤モデルを生成し、ロボット実験室を動かすクローズドループのAI実験プラットフォームを作成するよう指示。ライトDOE長官は**科学担当次官ダリオ・ギルをこのイニシアチブの指導者に任命**。

ゴール

世界最高のスーパーコンピュータ、実験施設、AIシステム、あらゆる主要な科学分野の独自のデータセットを統合したプラットフォームを開発し、**10年以内に米国の研究とイノベーションの生産性と影響力を倍増**させる

科学安全保障プラットフォーム (American Science Security Platform) の構築・運用

DOE国立研究所の世界最高のスーパーコンピュータ等の高性能計算資源、AIエージェントを含むAIシステム、計算ツール、ドメイン特有の基盤モデル、データセットへのアクセス、自律実験・製造を可能にする実験ツール等の機能を統合

90日以内：利用可能な連邦政府および潜在的な産業パートナーのリソース(計算、ストレージ、ネットワーク)を特定

120日以内：初期データセットの特定とプラットフォームへの導入計画策定

240日以内：AI主導の実験・製造を行うための施設の能力を評価

270日以内：少なくとも1課題でプラットフォームの初期運用能力を評価

国家科学技術課題の特定

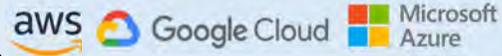
- ミッションを通じて対応可能と評価され、『国家科学技術覚書(9月23日)』に沿った**優先領域にまたがる国家的に重要な科学技術課題**について、**少なくとも20件のリストを、DOE長官は60日以内に科技担当大統領補佐官に提出**する。
- 2025年の初期リストには、「**先端製造**」、「**バイオテクノロジー**」、「**重要材料**」、「**核分裂および核融合エネルギー**」、「**量子情報科学**」、「**半導体およびマイクロエレクトロニクス**」を含む。
- リストの提出から**30日以内**に、科技担当大統領補佐官は提案されたリストを審査し、関係省庁と協力して、ミッションが対処すべき国家科学技術課題のリストを調整。リストは毎年見直し。



AI活用の促進

■STRIDESイニシアティブ

(Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability)



商用クラウドサービスの利用に伴う経済的・手続き的な障壁を減らし、生命医学研究を加速させることを目的に、NIH研究者・NIH支援研究者に、**主要商用クラウドの割引料金**のほか、**専門家コンサルティング**や、**高性能計算資源(HPC)**や**大規模データストレージへのアクセス**を提供

■NIH Cloud Labによる実践トレーニング

研究者のスキル向上支援のため、研究者が隔離された安全な環境で、AIやバイオインフォマティクスのチュートリアルを学習したり、概念実証の構築、コスト見積もり等を行うことができる**最大\$500のクレジット(最大90日間)を無償提供**(これまでの利用者は1,000人以上)

■安全なコード管理 NIH GitHub Resource Center

NIH所内 (Intramural) の職員 (研究者・開発者等) のみを対象に、研究者や開発者がコードを安全に保存・管理できるよう、高度なデータ損失防止機能やアクセス制御機能など、特別なセキュリティ機能を整備。また、個人のGitHubアカウントでは得られない多くのNIH特有の機能やサポートも提供。

■AI・クラウド環境の活用事例

- ・TOPMed (Trans-Omics for Precision Medicine)
STRIDESを通じた大規模な全ゲノムシーケンスデータのクラウドへの移行により大人数での同時アクセスを可能にし、共同研究を加速
- ・Advanced Imaging and Microscopy (AIM) Resource
ローカル環境のインフラ容量の制約を、クラウドの拡張性で補うことで、データ分析の高速化や複数の並行研究が可能に。

■AI活用促進の取組

- ・AIM-AHEAD 電子カルテを含む大規模データを基盤にAI/MLの研究・基盤整備・人材育成を推進し、健康格差の是正と医療革新を目指す全米規模のイニシアチブ
- ・Bridge2AI AIに適した高品質なデータの整備や、データの使い方や評価方法の共通ルール、解析ツールの提供、AI人材の育成体制の構築等を推進

リスクへの対応(研究の公正性の確保、データ管理)

■助成金申請におけるAI利用の制限(2025年9月～)

AIツールを用いて大量の申請が行われる懸念に対応するため、2025年9月以降、AIで実質的に作成された申請は審査対象外に。また、PIが1年間に提出できる申請数は最大6件に制限。

■査読(ピアレビュー)プロセスでのAI利用の使用禁止

未公開の研究アイデアの漏洩を防ぐため、査読プロセスにおける生成AIの使用は禁止。

■制限データ(ヒトゲノムデータ等)を使用する際の制限

●オープンなAIモデルへの入力制限

●開発したAIモデルの共有制限

- ✓ プライバシーの懸念がない一般的なデータを用いて開発されたAIツールやソフトウェアは、NIHの**オープンサイエンス原則**に基づき、無料でオープンな形式で共有することが原則。
- ✓ 他方、**ヒトのゲノムデータなどのアクセス制限データを学習させたAIモデルのパラメータ**は「データ派生物」と定義され、漏洩リスクを防ぐために、**プロジェクト終了後の保持や第三者への共有を、現在、一時的に禁止**

■プライバシー強化技術(PETs)等の検討

イノベーションを促進しつつ、制限データ漏洩リスクを軽減するための、プライバシー強化技術 (Privacy Enhancing Technologies) やその他の方策について、研究コミュニティや一般に向けて広く情報提供依頼 (RFI) を発出

英国 AI for Science Strategy (2025.11.21)

英政府は、AI for Scienceを加速するための**具体の15項目**を示し、「AI機会行動計画(AI Opportunities Action Plan, 2025年1月)」における2026～30年の政府投資 £ 20億のうち**最大 £ 1.37億**を充てると発表。また**優先 5 分野**を示すとともに、官民で挑む野心的な目標(**ミッション**)も掲げた。

【目的 (Objective)】

- AI を活用した科学 (AI-driven science) の最先端を切り開く能力を高める ⇒Action 1 & 2
 - 英国の科学分野におけるリーダーシップの地位を維持することを確実にする ⇒Pillar1～3の下に、Action 3～14
- }+ ミッション (Action 15)

【優先 5 分野】

英国が既に強みを持ち、AIにより大きなインパクトが期待される「**エンジニアリング・バイオロジー**」、「**核融合**」、「**材料科学**」、「**医療研究**」、「**量子技術**」(※「産業戦略(Modern Industrial Strategy)」とも連携)

■ AI駆動科学

- (Action 1) AI駆動科学を加速。Sovereign AI Unitが「**自律型ラボの開発・拡大**」について、またARIAが「AI scientist」について公募を実施。
また、責任ある安全なAI活用について、**バイオセキュリティへの影響などドメイン特有の課題にも対処**。
- (Action 2) メタ科学 (metascience) の研究を支援。AIが科学研究そのものをどのように変えるか、UK Metascience Unitと連携して、探求。

■ データ (Pillar 1)

- (Action 3) UKRIが資金提供した研究によるデータを可能な限り保存・キュレーションし、FAIR 原則に準拠させる。2026年にデータポリシー更新予定。
- (Action 4) DSITは、**高価値データセットを特定し、整備**。Renaissance Philanthropyと連携し、まず優先5分野のデータを精査・収集を開始。
- (Action 5) UKRIは、ネガティブ実験データも含む**ダーク・データ (dark data)の収集パイロットプログラム**の開始。
- (Action 6) 高価値データセットの安全な活用のための**大規模なデータストレージインフラ**を、国のスーパーコンピューターの近くに整備。

■ 計算資源 (Compute) (Pillar 2)

- (Action 7) AI Research Resource(AIRR)を通じた公募を開始。**研究者向けにIsambard-AIやDawnスパコン上でのGPU時間を提供**。
小規模向け：最大10,000 GPU時間/3カ月間、ミッション重点プロジェクト(医薬品開発等)：最大1,400,000GPU時間(2週間程度)
- (Action 8) 英国内の研究機関を結ぶフェデレーテッド・コンピュート・クラスタ・ネットワークの構築。未使用/余剰の計算資源の活用可能性を追求。

■ 人材と文化 (People and Culture) (Pillar 3)

- (Action 9) 今後5年間、AI for Science博士課程プログラムの拡充等を通じて、**AIを研究に活用できる研究者を1,000人以上育成**する。
- (Action 10) 学際的なフェローシップ・プログラムを活用し、**トップ研究者にAI能力を身につけさせながら、コミュニティのスキルアップ**を図る。
- (Action 11) あらゆる分野やキャリア段階の研究者や技術専門職(RTP)のAIスキル向上のための**トレーニングプログラムの創設・提供**。
- (Action 12) 優先分野におけるAI for Scienceのブレークスルーを推進するための**学際的研究チームの構築** (UKRIによる「ハブ」型研究支援)。
- (Action 13) 研究技術専門職(Research Technical Professional RTP)の育成・確保への投資と、そのためのキャリアパスの整備。
- (Action 14) **AIモデルのコミュニティ主導のベンチマーク・評価の開発支援**。(例: 構造生物学のCritical Assessment of Structure Prediction)

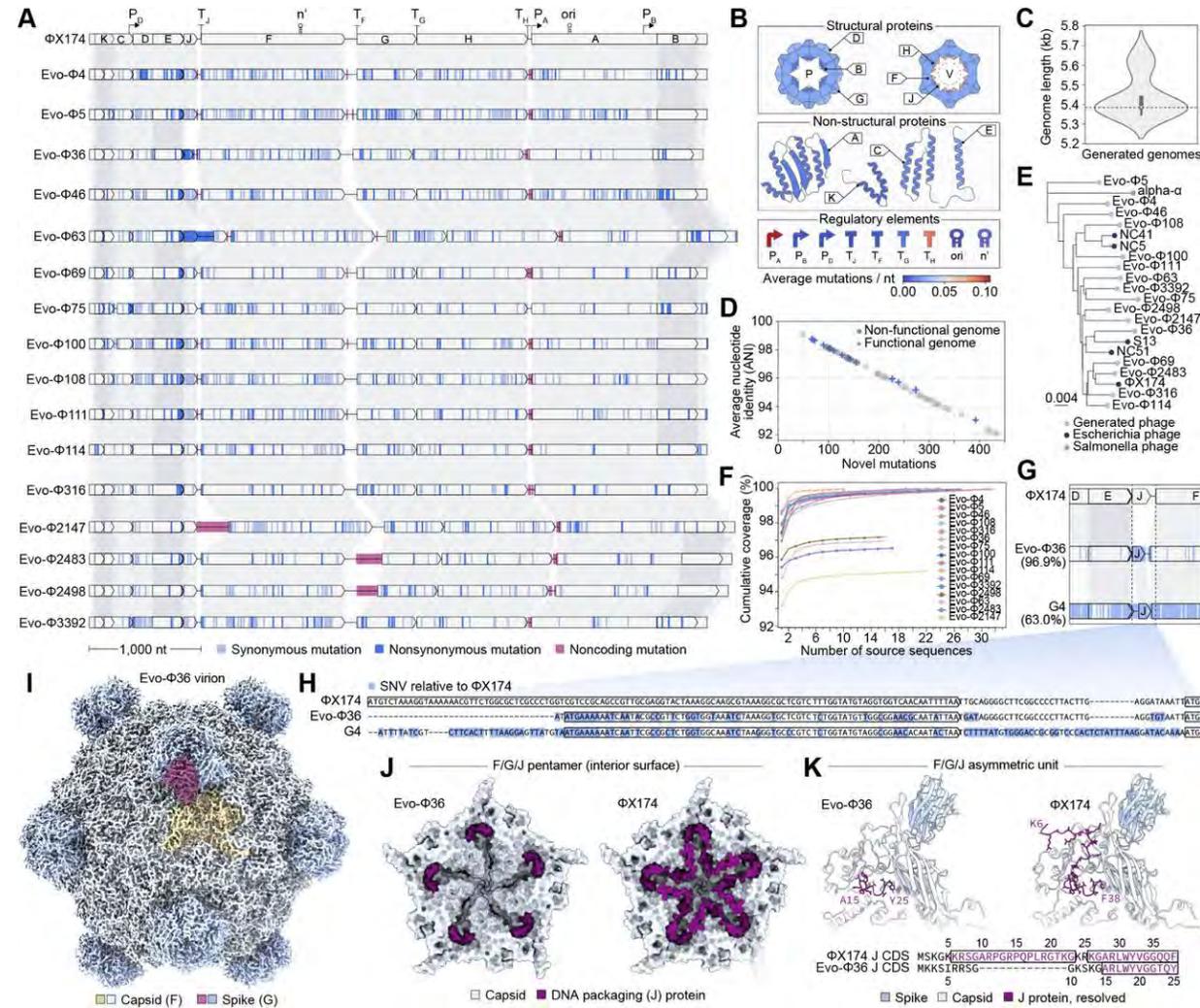
■ ミッション (Missions)

- ミッション 1** は、「**前臨床段階(develop trial-ready drugs)を2030年までに100日以内とできるように創薬を加速し、新治療薬の早期展開に貢献**」
- (Action 15) 追加のAI for Scienceミッションをいくつか選定し、**2026年に開始**。(GO-Scienceのライヴ・インスタンシング 等も活用、アカデミア等とも対話)

(参考)ゲノム言語モデルを用いたバクテリオファージの設計・生成(2025年9月)

2025年9月、米・Arc Institute(非営利研究機関)とスタンフォード大の研究チームが、Evo1及びEvo2モデルを用いて、抗生物質に耐性のある大腸菌株に感染するバクテリオファージを設計・生成したことを発表。

- Evo1およびEvo2は既に200万以上のファージゲノムを用いてトレーニングされているが、研究チームは、さらに「教師あり学習」で追加学習をおこなったうえで、5,386塩基、11個の遺伝子から構成されている、比較的小さなゲノムを持つ一本鎖DNAウイルスである ϕ X174をテンプレートとして、抗生物質に耐性のある大腸菌株に感染するようなゲノムを生成。
- 研究者らはAIの生成したゲノム情報に従ってDNAを合成し、それをファージとして増やした後、目的とする大腸菌に感染させ、それを死滅させることができるかどうか調べた。AIモデルが提案した302種類のゲノム設計のうち16種類が実際に機能し、大腸菌の成長を阻害したり死滅することを確認。

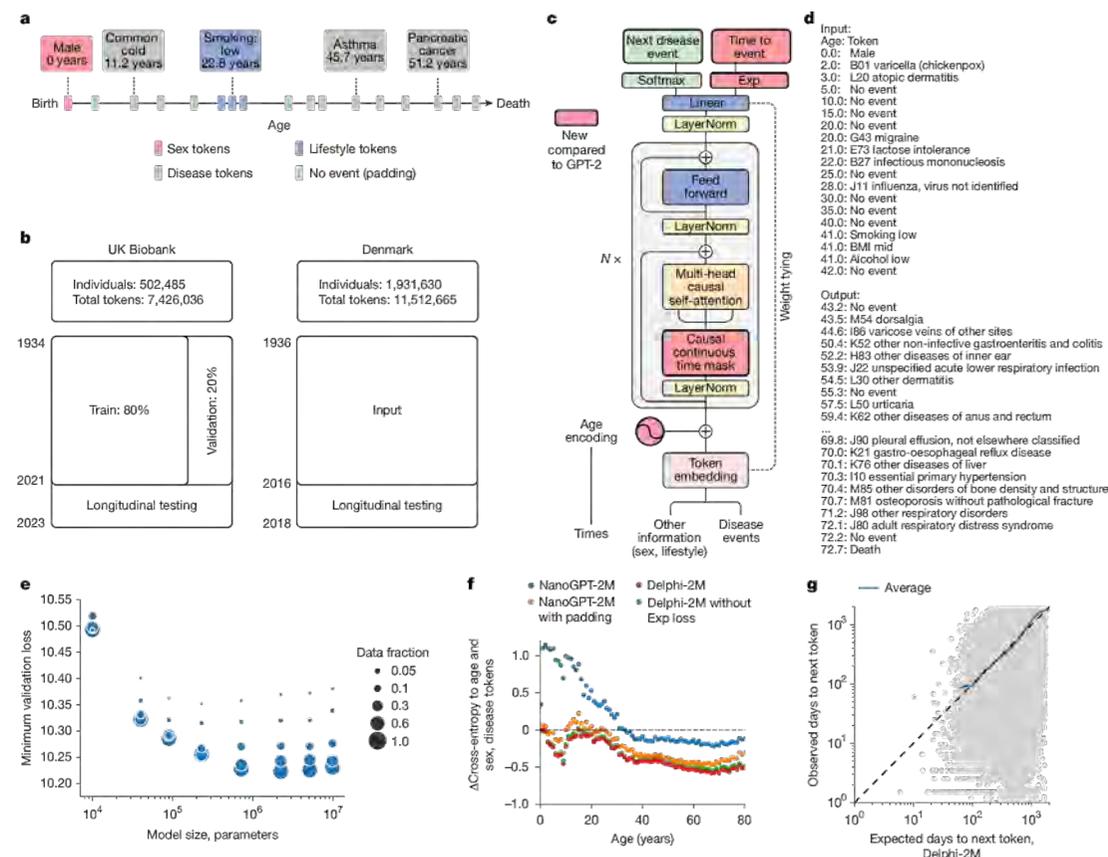


(資料) <https://www.biorxiv.org/content/10.1101/2025.09.12.675911v1>
<https://arcinstitute.org/news/hie-king-first-synthetic-phage>
<https://www.technologyreview.jp/s/369287/ai-designed-viruses-are-here-and-already-killing-bacteria/>

(参考) 疾患予測モデル Delphi-2M (2025年9月)

欧州分子生物学研究所(EMBL)、独がん研究センター(DKFZ)、コペンハーゲン大学の研究チームが、英国バイオバンク(UK Biobank)およびデンマークの匿名化されたデータを活用し、患者の記録(生活習慣要因や他の健康状態など)における特定の疾患発生時期の相対的なパターンを検出するAIモデル「Delphi-2M」を開発。

- このAIモデルは、GPT-2のアーキテクチャを拡張したモデルにUK Biobankの40万人分のデータを用いて訓練され、デンマークの190万人の患者データを用いて検証された。研究チームは**個人の病歴に基づき1,000種類以上の10年以上先の疾患発生確率を予測**できることを確認。
- その精度は、各ツールで予測対象となる疾患数をはるかに少ない既存ツールと同等かそれ以上であり、**さらに最大20年先までの将来の健康経路をシミュレーション**し、プライバシーを保護しつつ他のAIモデルの訓練に有用な合成データを生成することも可能なことを確認。
- このAIモデルは、厳格な倫理規則のもと匿名化された健康データを用いて訓練されており、英国のバイオバンク参加者はインフォームド・コンセントを行い、デンマークのデータはデンマーク国内に留まることを義務付ける国内規制に従ってアクセスされている。
- なお、今後より多様な健康情報を取り込むことで、個別化医療の向上にも貢献するとみられるが、著者らは、本モデルが学習データに内在するバイアスを反映している点、および追加検証なしに直接的な医療判断に用いるべきではない点を指摘している。



出典) Shmatko, A., Jung, A.W., Gaurav, K. et al. Learning the natural history of human disease with generative transformers. *Nature* (2025), 17 September 2025
<https://www.nature.com/articles/s41586-025-09529-3>
<https://www.natureasia.com/ja-jp/nature/pr-highlights/15346>
<https://www.embl.org/news/science-technology/ai-model-forecasts-disease-risk-decades-in-advance/>

厚生労働省の取組について

令和8年2月27日 第2回 医薬品開発協議会

ひと、暮らし、みらいのために



厚生労働省
Ministry of Health, Labour and Welfare

第1回 医薬品開発協議会の資料より

< 医薬品開発協議会 >

医薬品開発・製造プロセスの効率化に向けて

第1回 医薬品開発協議会
第1回 感染症協議会
令和7年6月2日

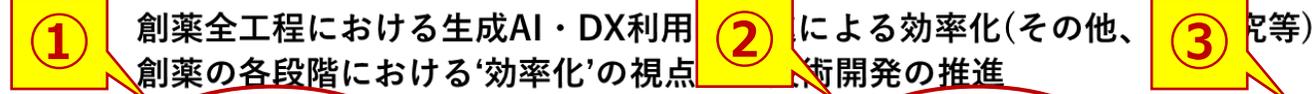
資料3-1

【背景・目的】

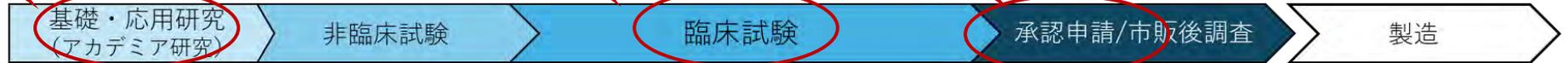
個別医薬品企業においては、開発・製造プロセスに対して効率化、生産性向上を目指した様々な取組がなされているが、医薬品の研究開発費や製造コストは近年顕著に高騰し（図1参照）、国の医療費の高騰の一因ともなっている。また、現在、医薬品開発の主体となっているベンチャーやアカデミア研究者にとっては、研究開発費の高騰はその活動を阻害する最大の要因となっている。

そこで、本年度の医薬品開発協議会では、「AI創薬」をはじめとした、医薬品開発・製造のさらなる効率化（生産性向上）に資する施策を重点的に議論していくこととする。技術革新の観点から大胆な効率化に挑戦し、技術自体の開発とプロセス全体の変革を目指す。

【医薬品開発・製造における効率化施策の全体像】（案）



1品目当たりの開発コストは2020年26億ドル（約4000億円）。2003年から145%増加。



AI
活用例

< DAIIA事業 >

- ✓ 候補物質の設計・探索・最適化
- ✓ 生物活性・毒性・動態予測

- ✓ 最適な治験デザイン提案
- ✓ Site selection/feasibility
- ✓ 患者リクルート

- ✓ 申請資料作成 (CSR・添付文書等)
- ✓ 安全性シグナル検知

- ✓ 製造工程開発
- ✓ 品質管理

その他
施策

- ✓ オルガノイドを用いた疾患モデル構築・標的同定

- ✓ 動物代替試験法の開発 (生体模倣システム(MPS))
- ✓ 分子シミュレーション (リード分子の結合親和性評価等)
- ✓ 自動化合成装置・ロボティクス (最適化サイクルの高速化)
- ✓ オルガノイドを用いたヒト組織での毒性・有効性試験

- ✓ 患者レジストリを活用した症例登録の強化 (特に希少疾患)
- ✓ DDC(Direct Data Capturing)によるモニタリング効率化
- ✓ DCT: ウェアラブルデバイスによるリアルタイム生体情報を用いた治験

- ✓ 新規モダリティ製品に対するガイドライン策定

- ✓ 連続生産 (製造の効率化・スケールアップ容易化)
- ✓ 3Dプリンティング
- ✓ GMP教育 (VRを用いた訓練)
- ✓ ロボットによる製造の自動化
- ✓ デジタルバリデーション (仮想環境での製造バリデーション)
- ✓ バイオ×デジタル製造基盤 (スケールアップの予測性向上等)

* Mullin, R., & Chemical Engineering News. (2020, November 11). Cost to develop new pharmaceutical drug now exceeds\$2.5B. Scientific American.

医薬品開発におけるAIの活用に関する厚生労働省の取組

① 基礎・応用研究段階における活用

「AMED 創薬支援推進事業（産学連携による創薬ターゲット予測・シーズ探索AIプラットフォーム開発）」にて、先行事業DAIIAで構築したAIを含む複数のAIを統合した創薬AIプラットフォームの構築を目指し、低分子のみならず、中分子・高分子にも適応モダリティを拡張する。

◆ 令和7年度 当初予算額 4.6億円、令和7年度 補正予算額 2.0億円、令和8年度 当初予算案 4.6億円

② 臨床試験段階における活用

「AMED 医療技術実用化総合促進事業」にて、AIの利活用が想定されるユースケース（研究参加者リクルートメント支援、各種文書の作成支援、治験データの分析・解析、リスク分析等、審査業務補助）について、少数の機関で試験運用を行い、実証フェーズへ向けて簡易的な手順書を作成する。

◆ 令和7年度 当初予算額 27億円の内数、令和7年度 補正予算額 22億円の内数

③ 承認申請段階における活用

「厚生労働科学研究費補助金（臨床研究等ICT基盤構築・人工知能実装研究事業）」にて、生成AIを活用し、医薬品・医療機器等の承認申請に関係する承認申請資料等を、効率的かつ高精度に作成するための技術基盤の構築、社会実装に向けた方向性の提言を行う。

◆ 令和7年度 当初予算額 0.3億円、令和8年度 当初予算案 0.3億円

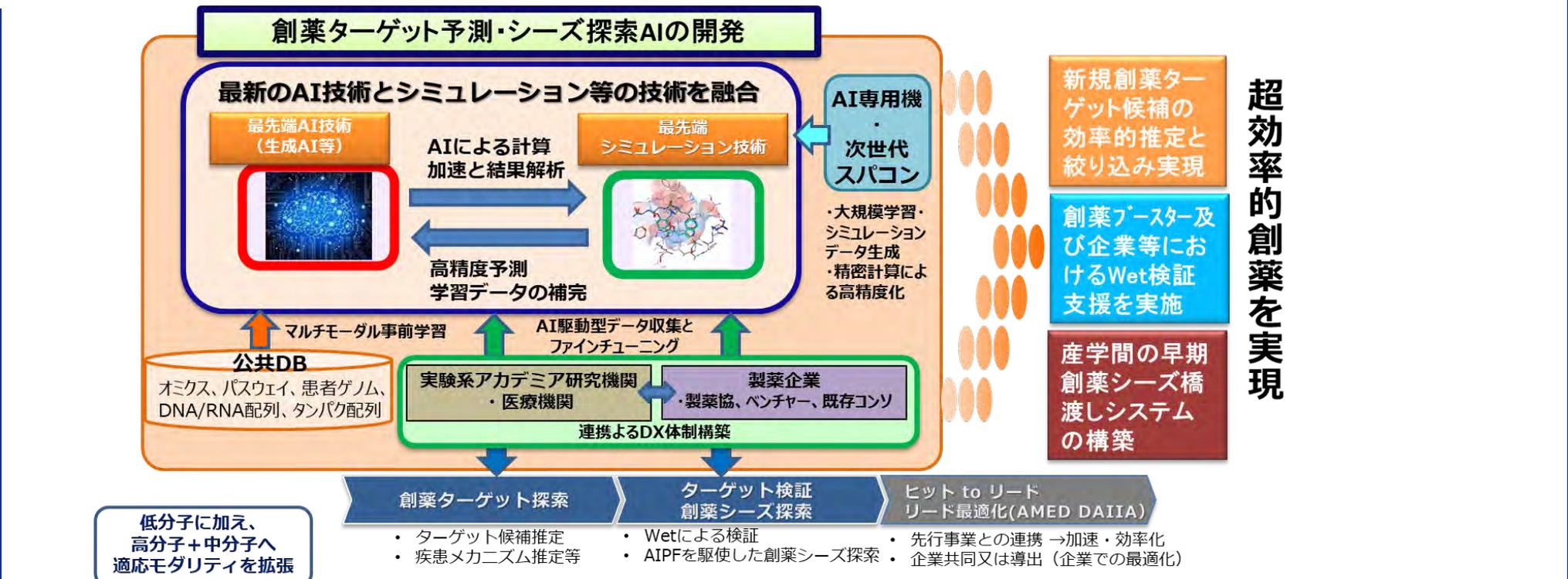
① 創薬支援推進事業 (産学連携による創薬ターゲット予測・シーズ探索AIプラットフォーム開発)

令和8年度当初予算案 4.6億円 (4.6億円) ※ ()内は前年度当初予算額 ※令和7年度補正予算 2.0億円

1 事業の目的

最新の高度なAI技術に基づく“創薬ターゲット予測・シーズ探索AI”を開発し、産学連携による効率的なデータ収集と予測精度の向上を図るとともに、創薬プロセスにおける複数のAIシステムを統合した創薬AIプラットフォームを構築し、創薬プロセスの効率化と創薬ターゲットの枯渇問題の克服を目指す。Wet検証やプラットフォーム構築を加速するため、要素技術の開発を強化する。

2 事業の概要・スキーム



創薬プロセスにおける複数のAIシステムを統合した創薬AIプラットフォームの構築を目指す。

3 実施主体等

◆ 補助先：国立研究開発法人日本医療研究開発機構 (AMED) ◆ 補助率：定額 ※AMEDにおいて公募により研究者・民間事業者等を選定

②

医療技術実用化総合促進事業 治験・臨床試験におけるAI利活用の推進に係る取組み

取組の概要

- 国際競争力のある治験・臨床試験実施体制を整備するため、治験DXに取り組むことは急務である。特にAI関連技術の利活用は治験・臨床試験の効率化に有用な手段であり、迅速かつ強力で推進する必要がある。そのため、今年度に新たにAIの取組を立ち上げる。
- 令和7年度は、AIの実装が想定される具体的なユースケースについて、少数の機関で試験運用を行う。

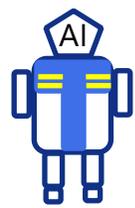
想定ユースケース

①研究参加者リクルート等	②各種文書の作成支援	③治験データ等の整理・分析	④その他
<ul style="list-style-type: none"> 電子カルテや診療記録等をAIが解析し、治験のフィージビリティ調査や研究参加者リクルートに活用 	<ul style="list-style-type: none"> 既存のプロトコルを参考に、AIが新たなドラフトの自動生成 症例報告書や同意説明文書などの各種文書の校正、ドラフト作成 英語で書かれた資料の翻訳 	<ul style="list-style-type: none"> 治験で収集した症例データのクリーニング、構造化 AIによる統計解析支援 安全性情報のスクリーニング、リスク検出 	<ul style="list-style-type: none"> 過去に実施された倫理審査委員会を題材に、生成AIを用いた倫理審査の実施可能性を検討



試験運用を通して、次年度以降の実証フェーズへの簡易的な手順書を作成する。
あわせて、院内システム導入時や運用時の技術的・組織的な課題について抽出し分析する。

実施期間等



実施期間：2025年12月～2026年3月
採択数：6施設
補助額：1施設あたり300万～400万

今後のタイムライン想定

※2026年度以降は予算獲得状況による

	2025年	2026年	2027年
計画・準備	→ 手順書作成・課題抽出		
実証		→ 複数機関への適用	
評価・展開			→ 運用ガイドライン策定

②【臨床研究中核病院に対する国際水準の治験・臨床試験対応能力の強化】 医療技術実用化総合促進事業

令和7年度補正予算 22億円

① 施策の目的

医療法に位置づけられている臨床研究中核病院において、その臨床研究基盤及びネットワーク機能を活用した臨床研究中核病院間の連携、各臨床研究中核病院の特色を生かした機能強化を推進するとともに、臨床研究・治験実施に係るノウハウを臨床研究中核病院外に共有・展開することで、日本全体の臨床研究基盤を強化し、日本発の革新的医療シーズ等をいち早く実用化に繋げ国民へ還元する取組みを推進していくことを目的とする。

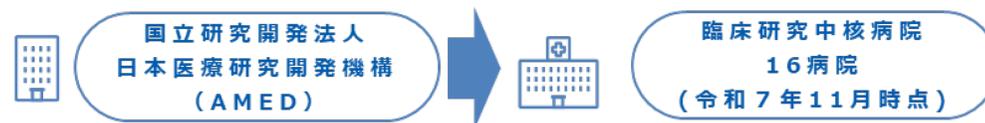
② 対策の柱との関係

I			II					III	
1	2	3	1	2	3	4	5	1	2
			○						

③ 施策の概要

治験・臨床研究領域において、米国を筆頭に生成AIの利活用等DXが急激に進行しており、医療機関においても効率的な治験・臨床試験の実施に際して利活用の実例が現れた。我が国も国際的な潮流に遅れを取らないことが緊急の課題である。能力ある臨床研究中核病院に対して生成AIの利活用パイロット事業を始めとする治験・臨床試験DXの推進、国際ARO機能の強化、革新的医薬品等開発に取り組むベンチャーの臨床試験実施への取組を強化するため、本事業で実施中のプログラムに以下の項目を加える。

④ 施策のスキーム図、実施要件(対象、補助率等)等



(1) 国際共同臨床研究実施推進プログラム

- 臨床研究中核病院の臨床研究支援部門は、海外製薬企業やアカデミアとの対応可能な人材配置及び体制構築が必須となりつつある。そのため、1. 国際共同試験の企画・立案、2. 国際共同試験の調整事務局(海外との交渉・調整、資料翻訳、規制対応)、3. 自施設発シーズ等の国際展開・海外有望シーズ等の誘致 4. 国際教育等に対応できる体制を構築する。

(2) 先進的臨床研究実施推進プログラム

- 生成AIを用い、治験・臨床試験分野の効率化のためのパイロット研究を実施、診療録と治験データの統合・eWSの普及等といった治験・臨床試験DXの推進等の先進的な取組について、臨床研究中核病院において試行的な導入を行い、導入にむけてのノウハウを集積し、臨床研究中核病院外に共有・展開することで、国内での普及推進につなげる。

(3) 医療系ベンチャー育成支援プログラム

- 革新的な医薬品等の開発に取り組むスタートアップ等を臨床研究中核病院が支援する体制を強化し、新たにスタートアップ等によるFIH試験を含む臨床試験の実施促進を図る。

◆補助先: 国立研究開発法人日本医療研究開発機構(AMED)

◆補助率: 定額

⑤ 成果イメージ(経済効果、雇用の下支え・創出効果、波及プロセスを含む)

臨床研究中核病院において、治験・臨床試験の企画立案、審査プロセス等において、生成AIを始めとするDXに対応することを通じて、治験・臨床試験の効率化を促進するとともに、グローバル対応に向けた体制整備を行い、国際水準の治験・臨床試験の実施を推進する。

③ 令和7年度厚生労働科学研究費補助金（臨床研究等ICT基盤構築・人工知能実装研究事業） 生成AIを活用した薬事承認申請・審査関連文書作成の推進のための研究

- 生成AI技術を活用し、医薬品、医療機器等の承認申請に関係する承認申請資料等を効率的かつ高精度に作成するための技術基盤の構築、社会実装に向けた方向性の提言を行う。

<目的>

近年では大手製薬企業のみならず、アカデミア発の医師主導治験やスタートアップ企業による開発が活発化しており、総括報告書（CSR）や承認申請資料の作成を高い信頼性かつ短期間で行うことは大きな負担となっている。また、医薬品医療機器総合機構（PMDA）にとっても、提出文書の質が審査の迅速性・正確性に直結するため品質確保が重要であり、審査報告書の作成においても多くのリソースを費やしている。また、海外においては、米国食品医薬品局（FDA）において審査業務を効率化するために大規模言語モデルであるElsaが導入された。日本では国立がん研究センター中央病院（NCCH）が行った生成AIを活用したCSR自動作成のパイロット研究の結果、8割以上の出力が微修正以下で実用可能という成果を発表している。

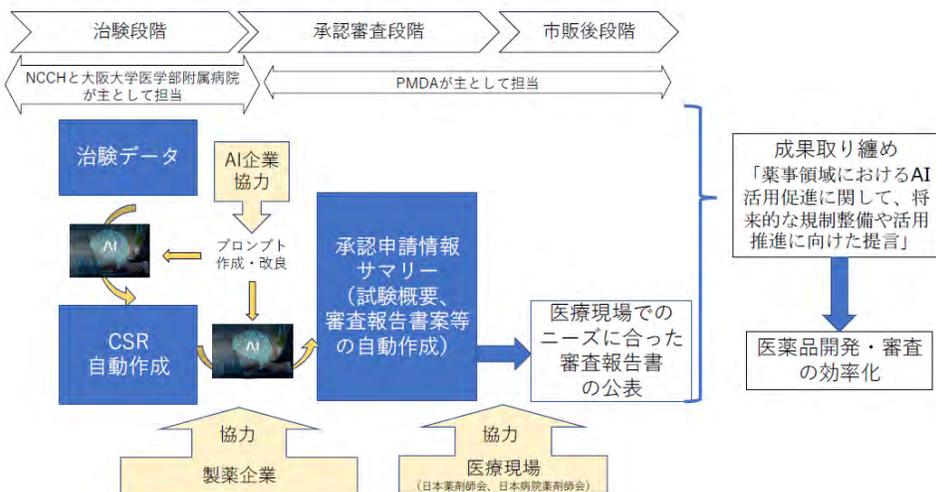
これらの背景を踏まえ本研究では、生成AI技術を活用し、CSRや医薬品・医療機器等の承認申請に関係する情報のサマリー（審査報告書を含む）を効率的かつ高精度に作成するための技術基盤の構築、社会実装に向けた方向性の提言を行うことを目的とする。

研究期間：令和7年10月～（3カ年計画の1年目）

研究代表者：中村 健一（国立研究開発法人 国立がん研究センター）

研究費：30,000千円（令和7年度）

研究の流れ図



<各年度の目標>

令和7年度：

パイロット研究とは別の医師主導治験を題材にCSRを自動作成し、さらなる精度向上を図る。PMDAでは、承認申請情報サマリーを文書生成AIにより作成し初期モデルを構築するとともに、文書生成AIを活用した研究環境の強化を図る。初期モデルの改善に向け、関係者の協力を得て検討事項を整理する。

令和8年度：

初年度とは異なる複数の医師主導治験を題材にCSRの自動作成に取り組むとともに、他機関の資料を用いたCSR自動作成も検討し、より汎用性のある技術基盤を策定する（企業治験への適用可能性の検討を含む）。承認申請情報サマリーの自動作成に必要なプロンプトの改良を関係者の協力も得ながら行う。あわせて海外でのAI活用状況の調査を行う。

令和9年度：

ベンダーの支援なしにNCCH内でCSR自動作成が実稼働できる体制を構築する。PMDAでは、承認申請情報サマリー作成のプロンプトのさらなる改善とともに、質の向上に資するよう留意点として整理する。これらを通じて、薬事領域におけるAI活用に関して、将来的な規制整備や活用推進に向けた提言を行う。

経済産業省の取組

令和8年2月27日

経済産業省 商務・サービスグループ

生物化学産業課

次世代治療・診断実現のための創薬基盤技術開発事業

商務・サービスグループ

生物化学産業課

令和7年度予算額 58億円（53億円）

事業の内容

事業目的

医療の課題として、患者の方々のQOL（Quality of Life）を向上させるとともに、治療の適正化による医療費増加の抑制を図る必要がある。こうした背景から、個人差を踏まえたより効能の高い治療を実現する「個別化医療」の推進に向けて、**（研）日本医療研究開発機構を通じた医療基盤の技術開発を行い、医療分野の産業発展に貢献する。**

事業概要

「個別化医療」を推進する技術開発として、新モダリティとして注目されるRNAを標的とした医薬品の創薬技術（令和3～7年度）、国際競争力のある次世代抗体医薬品の製造技術（令和3～7年度）、マイクロバイーム制御による次世代治療技術（令和3～8年度）、次世代送達技術（令和6～11年度）等の研究開発を進めるとともに、バイオ医薬品の技術基盤の確立にもつなげる。【補助率：定額（10／10）、一部事業は2／3で実施】

事業スキーム（対象者、対象行為、補助率等）



※一部は2/3委託で実施

成果目標・事業期間

平成27年度から令和11年度までの15年間の事業であり、バイオ医薬品の国内製造技術基盤の確立を目指す。具体的には、

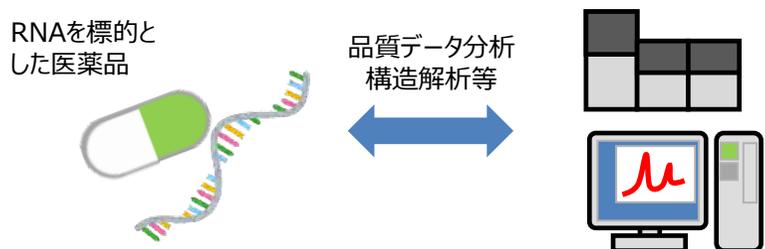
○令和12年度（2030年度）までに事業成果である各種技術を活用した我が国発の核酸標的医薬品シーズの前臨床試験の件数3件

○令和12年度（2030年度）までに事業成果である各種技術を活用した我が国発の抗体医薬品シーズの前臨床試験の件数5件（累計）等を目指す。

① RNA標的創薬技術開発

RNAを標的とした国産医薬品の実用化を促進するための基盤構築

- ・核酸医薬品の製造・精製・分析技術から規制まで一貫した基盤技術開発
- ・創薬標的となる核酸の立体構造・薬物ドッキングの解析・予測技術開発 等



② 国際競争力のある次世代抗体医薬品製造技術開発

次世代抗体医薬品を含む国産抗体関連技術の幅広い実用化推進

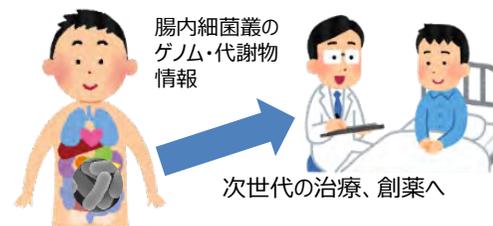
- ・次世代抗体医薬品等の製造基盤技術開発
- ・国産抗体生産細胞の商用化に必要な基盤技術の開発 等



③ マイクロバイオーム制御による次世代創薬技術の開発

マイクロバイオーム制御による医薬品の奏効率向上等、新たな創薬を可能とするマイクロバイオーム創薬エコシステムの構築

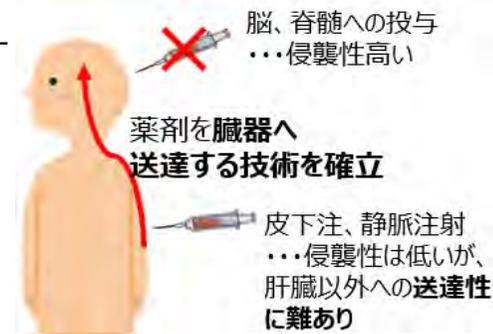
- ・マイクロバイオームの大量培養、安定製剤化技術開発
- ・マイクロバイオーム制御物質の探索・スクリーニング技術開発
- ・マイクロバイオーム医薬品の評価系の開発 等



④ 次世代送達技術開発

新規の（肝臓以外を標的とした）薬剤送達技術の創出と、それを用いた医薬品研究開発の推進

- ・有効成分・薬剤送達技術・リンカーといった要素技術の磨き上げ
- ・薬剤送達の評価技術の開発
- ・GMP準拠の製造実証と非臨床試験の実施
- ・実用化を見据えた知財戦略 等



①RNA標的創薬技術開発

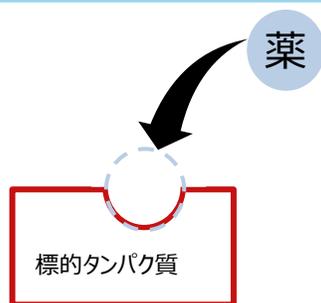
(AI活用に関連する公募研究開発課題の抜粋)

公募研究開発課題「**標的RNAの機能解析・構造解析基盤技術開発**」

- ・ 医薬品開発において**創薬標的の枯渇が課題となる中、創薬標的としてRNAが注目**されている。RNAの機能を制御する代表的な医薬品モダリティは核酸医薬であり、核酸医薬は、低分子医薬、抗体医薬に次ぐ第3のモダリティとなっている。
- ・ RNAの多くはタンパク質をコードしないnon-coding RNAであり、未だその多くは機能が解明されていないところ、その機能を予測する技術の確立に期待が寄せられている。また、**RNAの立体構造の情報は、RNAを制御する化合物を設計する上で非常に有用であるが、タンパク質と比較してRNAの構造情報は圧倒的に不足**しており、解析技術のバラエティも少ない状況である。
- ・ 本研究開発課題ではnon-coding RNAの機能を解析し、病態との関係を明らかにするとともに、解析手法の高度化により、**RNA標的創薬においても構造ベースの薬物設計を可能に**することを目指す。バイオインフォマティクス手法を用いたRNAデータベース等からのデータマイニング、リガンドの探索と相互作用解析、互換性を考慮したデータベース設計と構築、リガンドの複合体構造のデータベース化等を行い、AI 解析も活用し、創薬に活用できる技術を構築する。

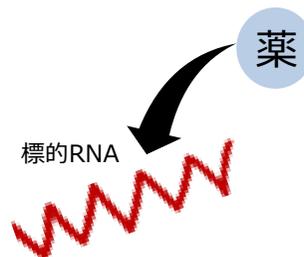
通常のタンパク質標的創薬

○タンパク質の立体構造データは十分蓄積



RNA標的創薬

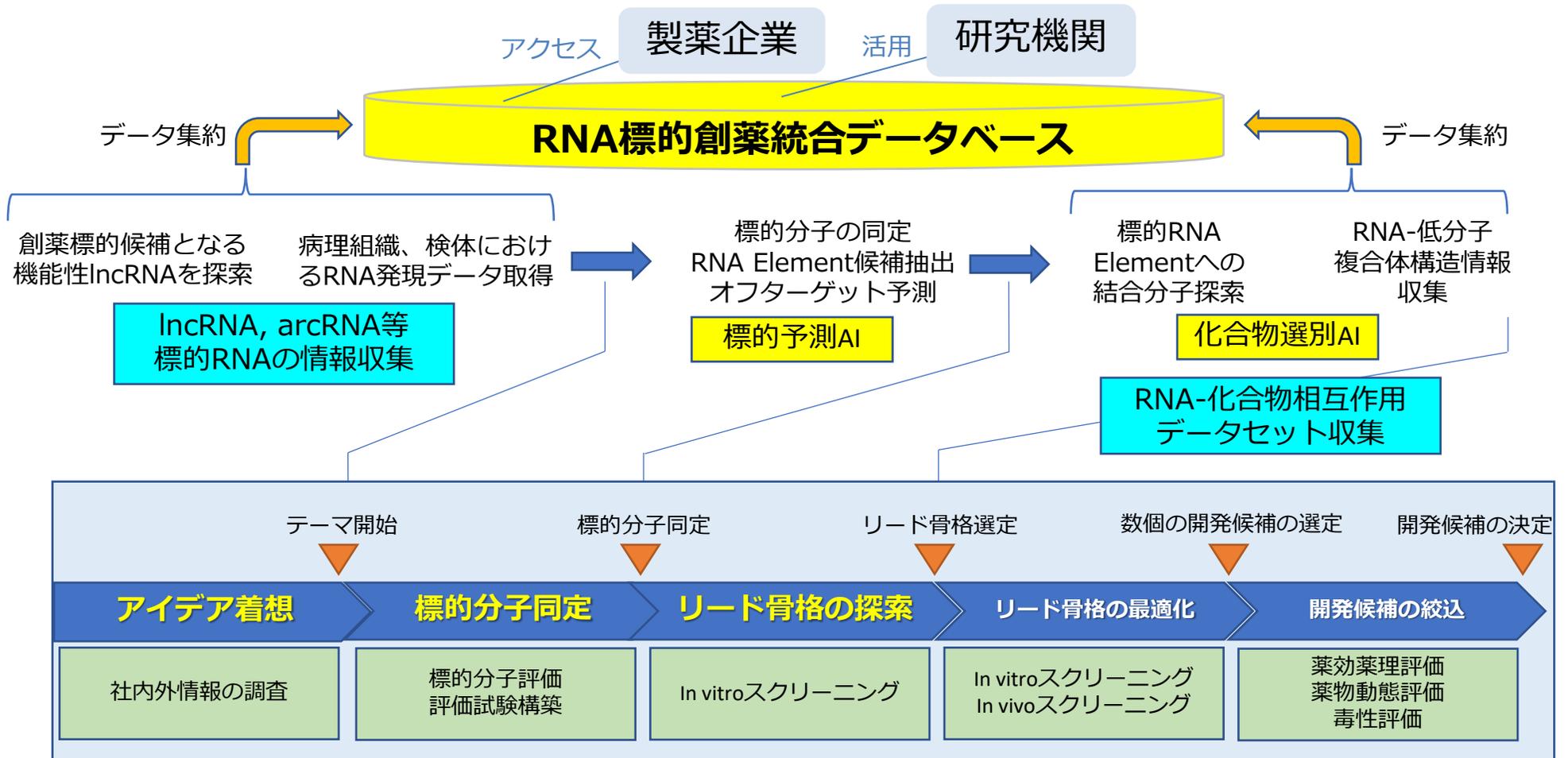
×RNAの立体構造データは蓄積が不十分であり、予測困難



① RNA標的創薬技術開発

(機能解析に基づくRNA標的創薬のための統合DBとAIシステムの構築 大阪大学)

- 病理組織のRNA発現解析情報、RNAと薬物の相互作用情報等を収集し、DBを構築
- 当該DBを利用して、RNA創薬標的の予測AI、標的RNAに結合する化合物選別AIを構築

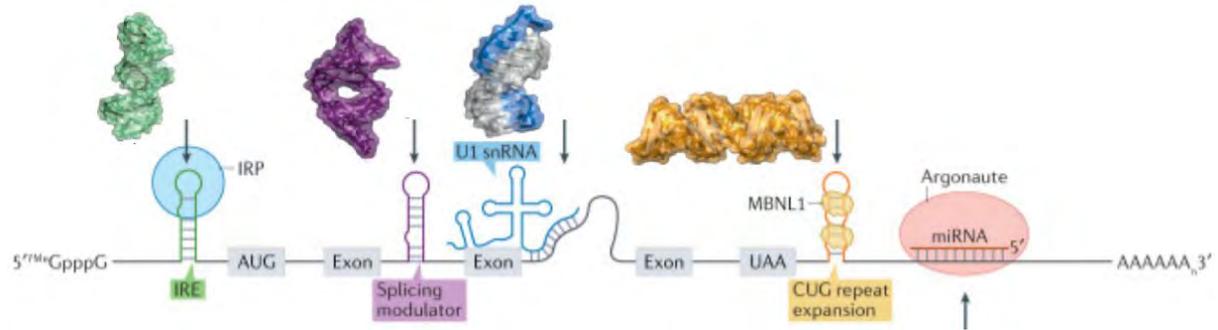
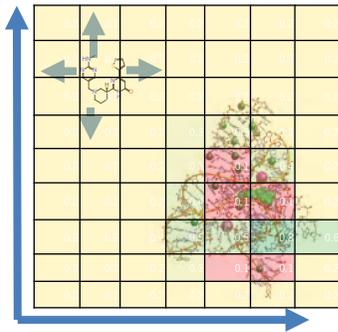


① RNA標的創薬技術開発

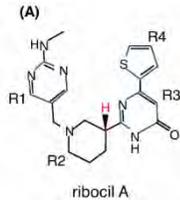
(RNA標的創薬に資するRNAおよびその複合体の機能・構造解析基盤技術の開発
 バイオ情報産業化コンソーシアム)

- 補助的にAIを利用した計算モデルにより、効率的・網羅的にRNAと薬物の相互作用を計算
- 当該計算モデルにより、RNA構造を標的とする薬物の分子設計を支援

CSD-mD-VcMD法
 (解析計算手法)



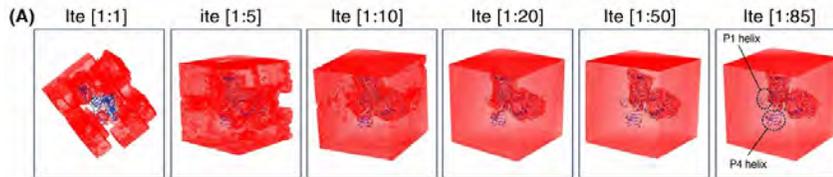
創薬標的となり得るmRNAの構造



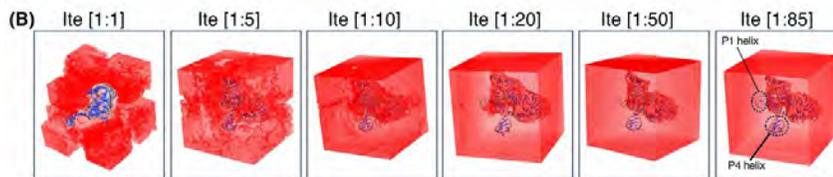
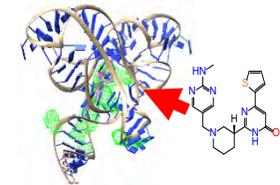
Ribocil A



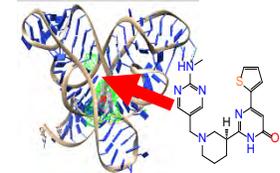
Ribocil B



不活性な薬物は不安定
 実験と一致



薬効の強い薬剤は安定、
 実験と一致



実験と計算を併用して開発

① RNA標的創薬技術開発

(RNA標的創薬に資するRNAおよびその複合体の機能・構造解析基盤技術の開発 バイオ情報産業化コンソーシアム)

- 本事業で収集したデータ及び開発ソフトを分子シミュレーションソフト (myPresto) で公開予定

myPrestoの構成

myPrestoは、化合物の二次元構造を三次元構造への変換、タンパク質等のモデリング、タンパク質-薬物ドッキング、in silicoスクリーニング等のプログラム群で構成されています。



②国際競争力のある次世代抗体医薬品製造技術開発

(AI活用に関連する公募研究開発課題の抜粋)

公募研究開発課題「バイオ医薬品製造プロセスの先端的シミュレーション技術の開発」

- **抗体医薬品の製造技術について、欧米への極端な依存**を改善すべく、これまで国内企業及びアカデミアが一体となり、競争力のある国産の抗体製造プラットフォームの構築、高品質の抗体を高効率で産生するホスト細胞の樹立や高度生産技術の開発を行ってきた。しかしながら、次世代医薬品の製造技術及び抗体医薬品製造の国産技術の社会実装はさらに強化する必要がある。
- 抗体医薬品などのバイオ医薬品共通の製造技術について、コストダウン、高品質化、少量多品種の製造等のニーズや、緊急時における迅速な製造体制の構築及び供給スケールアップへの対応に対する要請が高まっている。このような背景のもと、**品質とコスト、生産性のメリットを最大化する最適な製造プロセスの開発の効率化と迅速化**が期待される。
- 本研究開発課題では、これら社会ニーズに対応可能な次世代抗体等のバイオ医薬品製造技術に資する国際競争力のある先端的な研究開発を行う。具体的には、**バイオロジクス分野のデジタル・トランスフォーメーションを推進し、デジタルバイオマニュファクチャリング実現に関するバイオ医薬品製造プロセスのシミュレーション技術の開発**を行う。

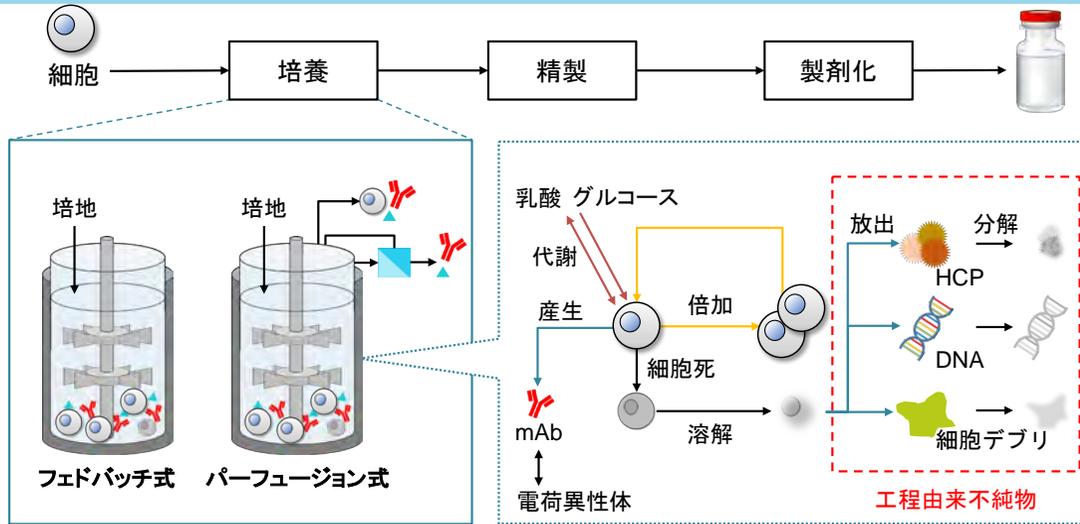


取得したデータからシミュレーション技術を開発し、効率化・迅速化・自動化へ



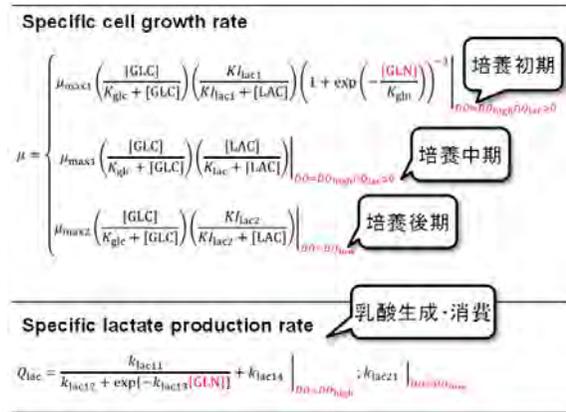
② 国際競争力のある次世代抗体医薬品製造技術開発 (バイオ医薬製造プロセスの開発・制御のためのデジタルプラットフォーム 東京大学)

- 培養プロセスについて、AIも活用しつつ物理的解釈可能な数理モデルを構築
- プロセスシミュレーションや品質・コスト評価に応用し、開発期間の短縮、製造コスト低減を目指す

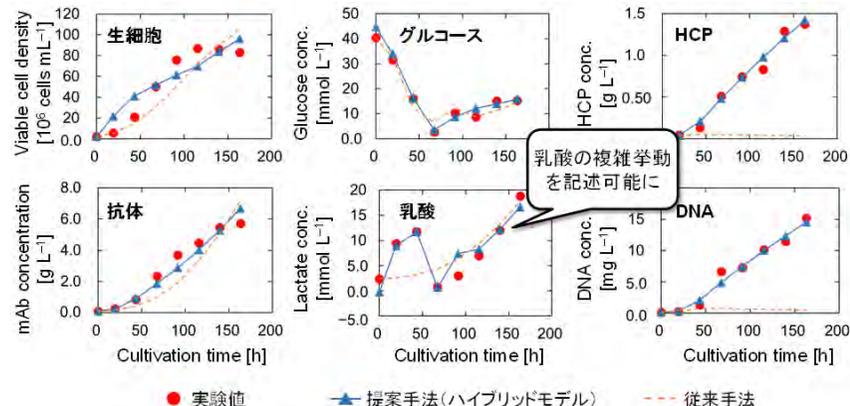


Rank	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅
1	F _p	Xv	μd	pH	T
2	T	DO	LAC	GLC	F _{glc}
3	V	P	μ	T	Xv
4	F _b	Time	Time	F _{glc}	μ
5	F	T	pH	μ	P

培養条件の寄与度分析



メカニスティックモデル



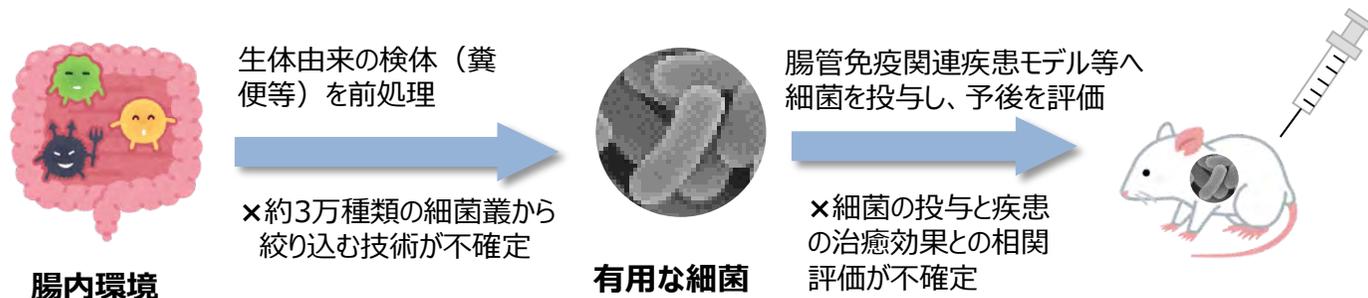
メカニスティックモデルとデータ駆動モデルのハイブリッドモデルによる予測

③マイクロバイオーム制御による次世代創薬技術の開発

(AI活用に関連する公募研究開発課題の抜粋)

公募研究開発課題「**腸管免疫関連疾患に対するMB制御技術応用による基盤技術の高度化**」

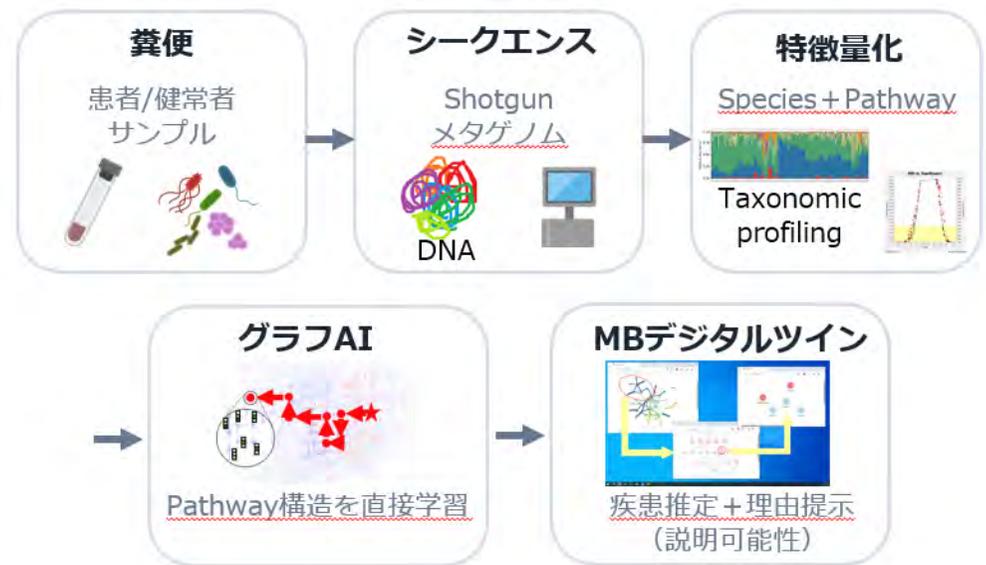
- 近年、**遺伝子解析技術の進歩により、マイクロバイオームのメタゲノム解析が大きく進展**し、さらに、**ノバイオーム技術の進歩によりMBが生体に与える影響や様々な疾患との関連性に関する知見が数多く報告され、MBに関する研究は著しく進歩している**。これらの研究の進展に基づく、多様な疾患に対する**MB制御による治療法の開発は新たな治療概念**になると考えられる。
- このような背景のもと、生菌製剤などの新たなモダリティで、MB制御による疾患治療技術の研究開発が国内外で急速に活発化している状況。しかし、我が国においては、生菌製剤特有の非臨床薬理及び薬物動態、毒性学的評価技術や製造・品質管理技術等は確立しておらず、国際競争力のある**MB制御医薬品の創出及び製造技術基盤の構築と幅広い実用化**を目指す必要がある。
- 本研究開発課題では、**MBを制御する有用腸内細菌の探索技術を軸としたMB制御菌製剤に関わる技術を高度化**する。具体的には、腸管免疫関連疾患患者からの検体採取や検体のMB解析（メタゲノム及びメタボローム解析）などを行い、創薬標的及びMB制御物質を見出す。



③ マイクロバイーム制御による次世代創薬技術の開発

(MBデジタルツイン技術を基盤とした腸管免疫関連疾患の制御法の開発 大阪公立大学)

- 糞便メタゲノムデータを教師データとしてAIに生体内経路を学習させ、MBデジタルツインを作製
- MBデジタルツインを用い、クローン病／パーキンソン病等のメタゲノム診断を目指す



- デジタルツイン：現実のデータをサイバー空間に写し、状態を再現・予測
- MBデジタルツイン：「腸内MBの双子」を糞便メタゲノムから構築し、疾患状態を推定

- 教師データ = 自施設で取得したクローン病／パーキンソン病の糞便メタゲノム
- 参照データ = 米・独・中など海外大規模集団



Roche ロシュグループ

中外製薬のAIやDXの創薬活用：抗体創薬を事例にとって

2026年2月27日 第2回医薬品開発協議会

研究本部モダリティ基盤研究部長

太田淳

中外製薬株式会社

まとめ

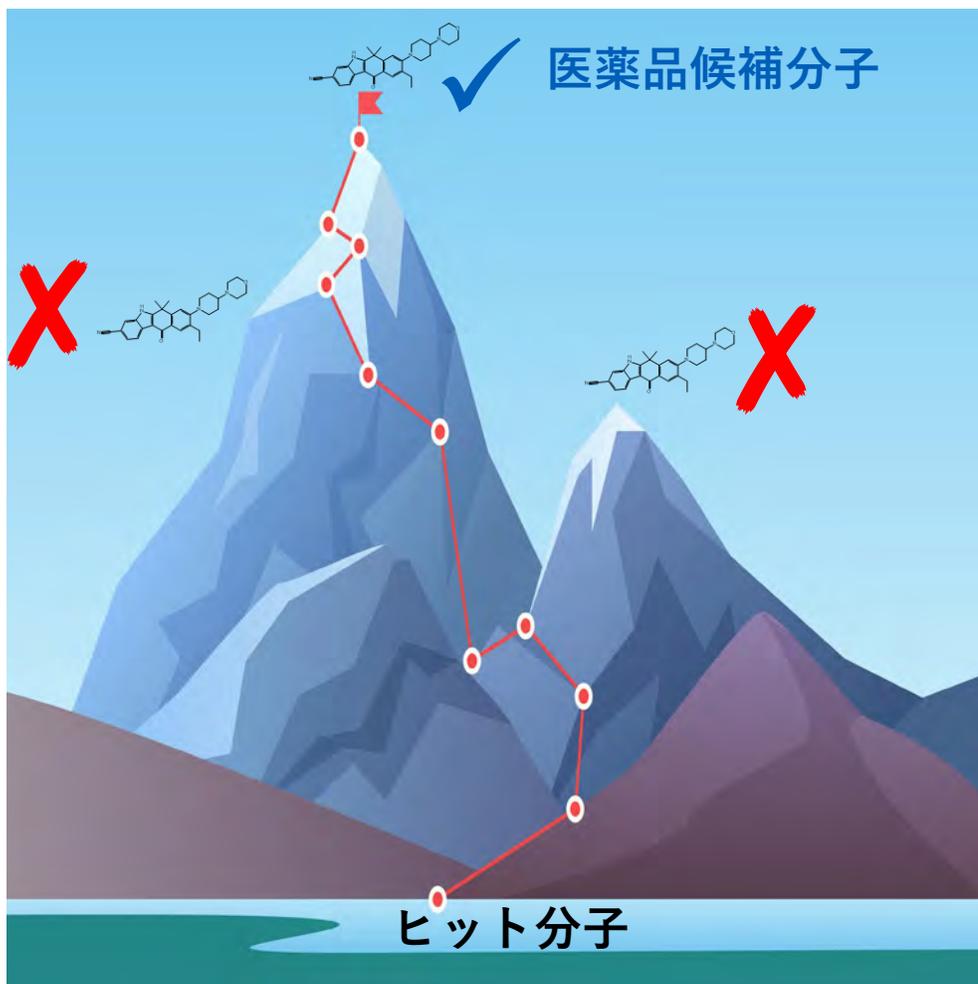
- 弊社抗体創薬は、他社では模倣が難しいような抗体工学（技術ドリブン創薬）や品質に立脚した（Quality-centric創薬）、比較的モノづくり難易度の高い創薬を実践している。
- その中で、①Lab Automationによる実験自動化、②データ成型と蓄積、③研究員のデジタルリテラシー向上、④AIによる分子デザイン、の4つの項目でモノづくりに関わるDXに挑戦してきた。
- 弊社としてはDXには、単なる効率化を超えたヒトの思考を拡張する補完的な技術としての役割（質的高度化）を期待している。

“Quality-Centric”の創薬

新薬候補分子の品質が中外製薬の最優先事項



Roche Roche Group



最高の医薬品候補分子を追求する強い社風

- 活性, 選択性, DMPK (Drug Metabolism and Pharmacokinetics (薬物代謝と薬物動態)), 安全性, 安定性といった観点において、最先端の技術で達成可能な最高品質の医薬品候補化合物を同定する
- スピードを優先することによって、医薬品候補分子の品質を損なうような創薬は行わない

“技術ドリブン”の創薬

- 中外製薬の創薬は「技術ドリブン」。特定疾患領域に限らず独自技術を医薬品に適用し、革新的な新薬の連続的創出を目指す
- 抗体エンジニアリング技術で世界をリード。低分子・中分子など多様なモダリティによる高い創薬力が強み

適切な創薬
ターゲットの選択
疾患原因分子

技術とターゲットの
マッチング

アンメットメディカルニーズへ対応

革新的医薬品創製技術の開発

低分子創薬

次世代低分子創薬の実現に向けて独自技術の確立と基盤プラットフォームの発展に注力



次世代抗体

国産初の抗体医薬品を創製した経験を活かし、独自の抗体エンジニアリング技術で新たな創薬につなげる



中分子創薬

低分子と抗体医薬の利点を併せ持つ中分子医薬品の創薬に10年以上前から投資し、独自の中分子創薬技術を確立



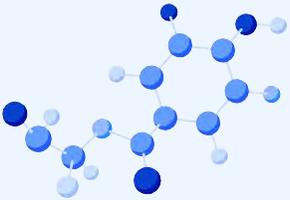
“技術ドリブン”の創薬

- 中外製薬の創薬は「技術ドリブン」。特定疾患領域に限らず独自技術を医薬品に適用し、革新的な新薬の連続的創出を目指す
- 抗体エンジニアリング技術で世界をリード。低分子・中分子など多様なモダリティによる高い創薬力が強み

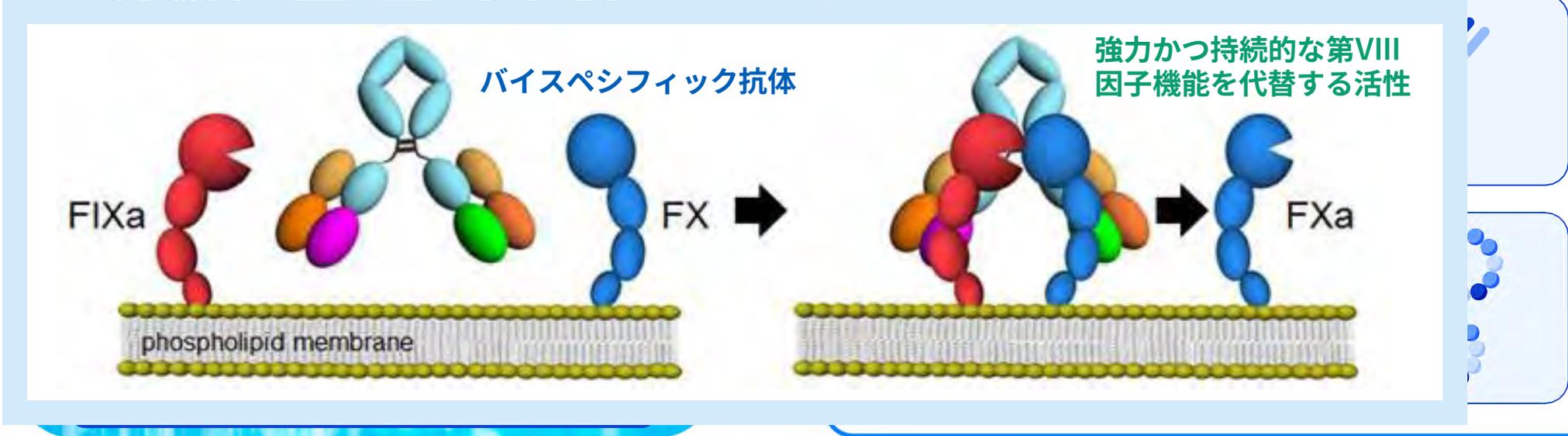
適切な創薬
ターゲットの選択
疾患原因分子

革新的医薬品創製技術の開発

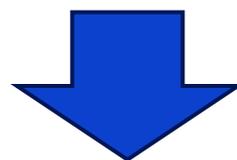
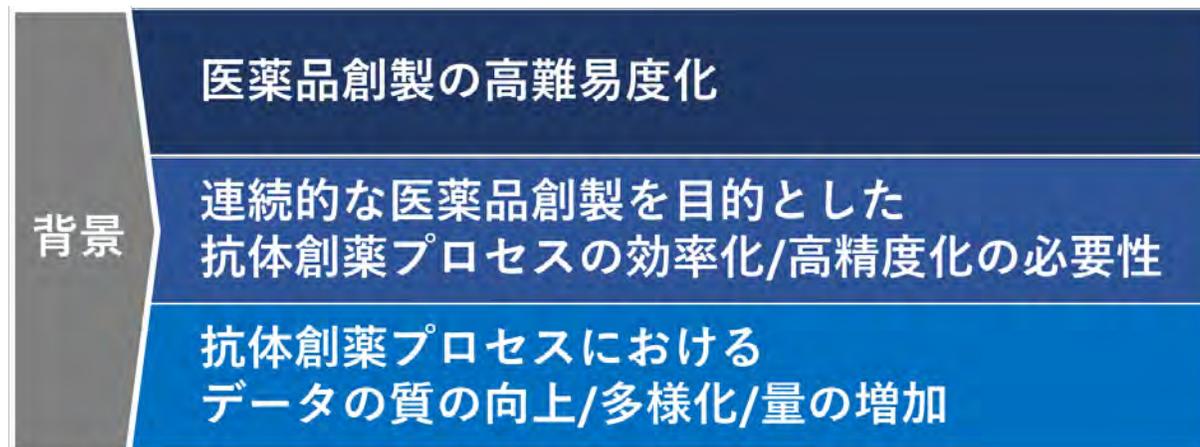
低分子創薬
次世代低分子創薬の実現に向けて独自技術の確立と基盤プラットフォームの発展に注力



技術とターゲットの



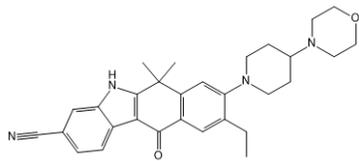
これらの高難易度創薬を進めるには生産性向上が不可欠



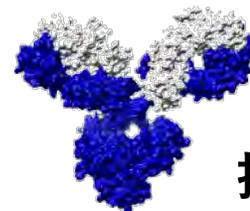
- **AIやデジタル/ロボティクス技術を使い創薬プラットフォームをさらに磨き続ける必要がある**
 - 革新的医薬品の連続創製
 - 開発期間の短縮、アウトプット倍増

複雑な操作をAutomation化し抗体創薬にも実装

低分子創薬



- より複雑な操作に対応する
- “中外の実験手技”を実現する

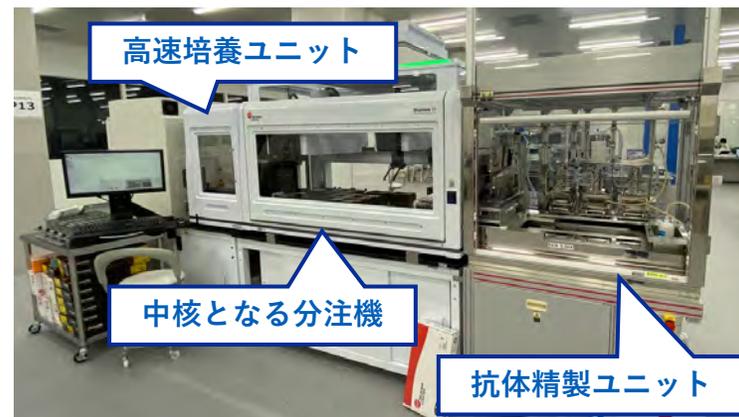


抗体創薬

例1：遺伝子クローニング自動化システム



例2：培養・抗体精製自動化システム



普段使っていない夜間の時間を利用する事で、従来5日かかっていた抗体遺伝子作製を3日に短縮

細胞培養(左側)と抗体精製(右側)の2つの実験を1台の分注機(中央)で実施。稼働率・投資効率を高めている

より柔軟な機器連携



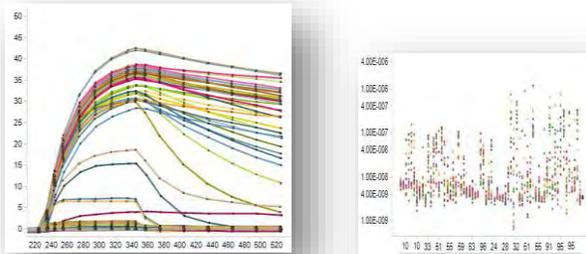
機器間の搬送

機器へのロード
・アンロード

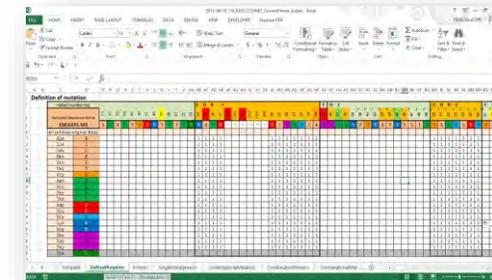
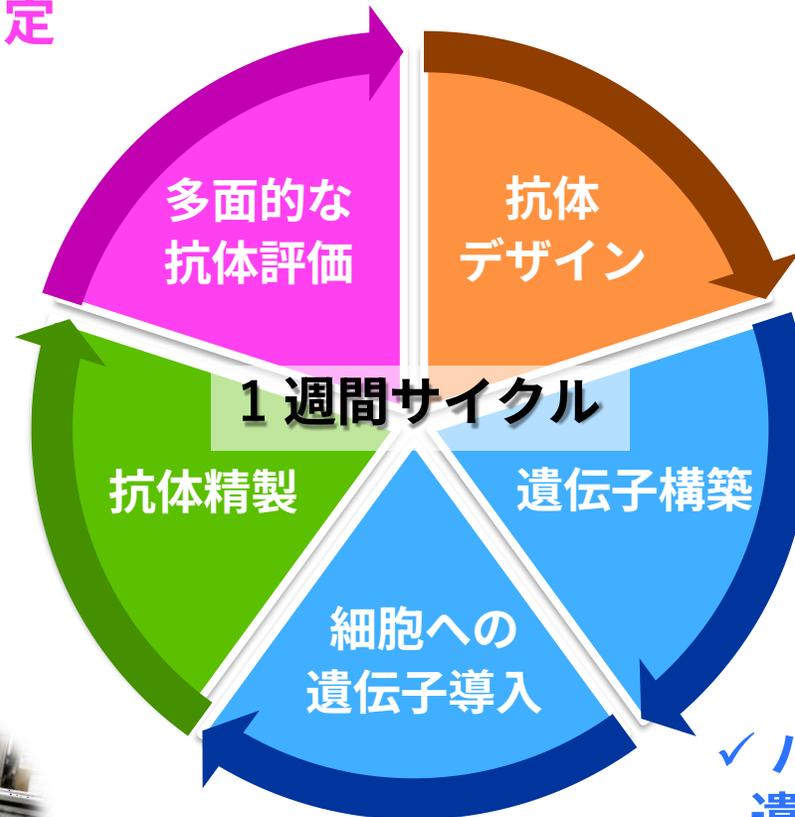
高スループットな抗体最適化プラットフォームを構築

1リード抗体につき約1300種類（抗原結合領域：約70か所 × 19アミノ酸）の抗体を作り多面的評価

- ✓ **ハイスループット結合活性測定**
約2000回/週
- ✓ **多面的な評価**（安定性、溶解性、免疫原性、非特異的結合など）



- ✓ **ハイスループット抗体精製**
約1500分子/日



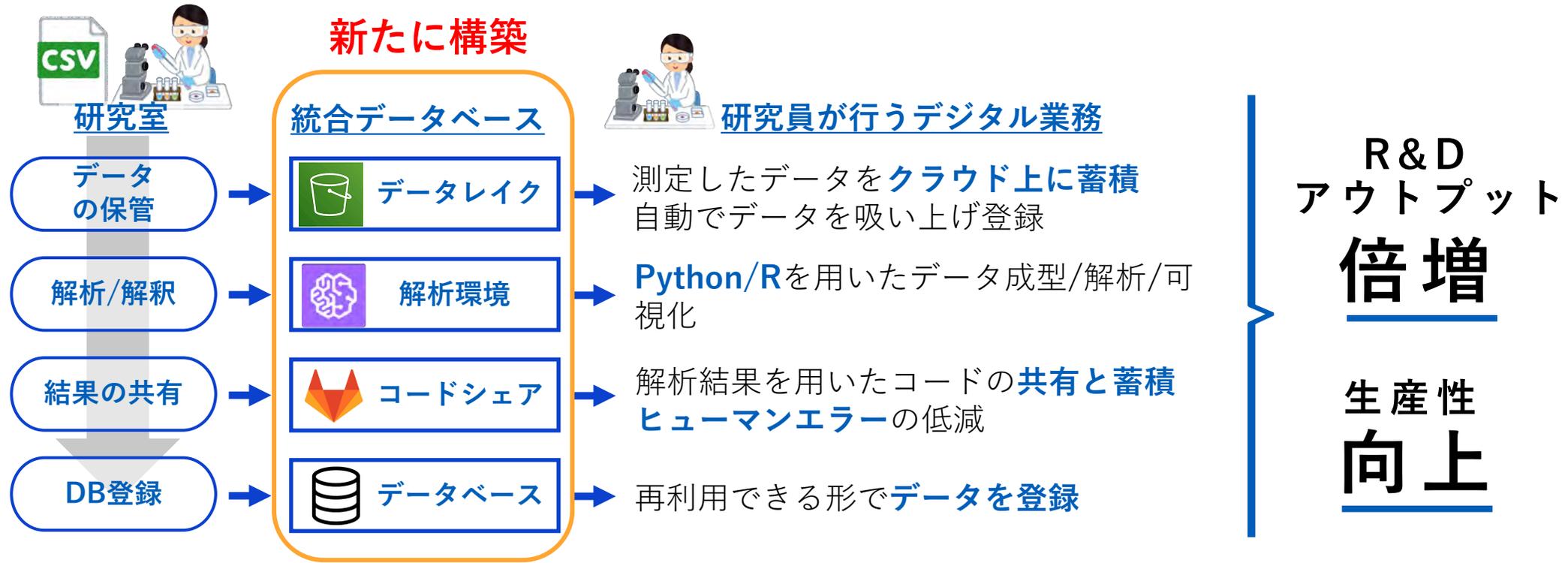
A screenshot of a software interface showing a data table with multiple columns and rows, likely representing experimental results or antibody characteristics.



- ✓ **ハイスループットな**
遺伝子構築と遺伝子導入
約3000分子/週

出典：中外製薬

大量のデータを効率よく利活用するデータベースを整備



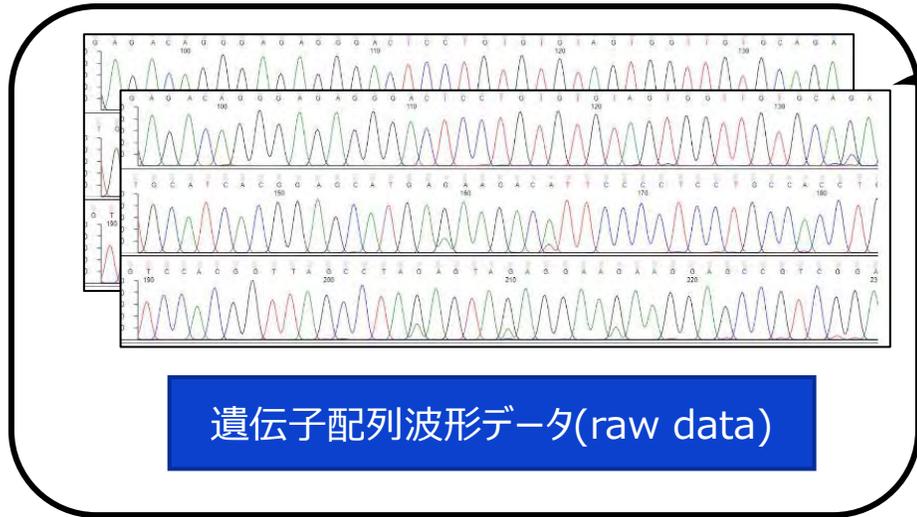
統合データベースを新たに構築。研究員はプログラミングを駆使してデータを扱うようになる。

各研究員がプログラミングで自身の研究を効率化

例) 抗体配列解析の自動化



step1: 夜間にメール受信を感知し波形データを自動で読み込む



step2: DNA配列解析システムで解析を自動実行



CHANGEシステム(中外)
遺伝子波形データを解析し目的配列と照合～齟齬を判定するシステム

step3: 解析結果メール自動送信



朝までに完了

Judgement	DNA	Sum	Ins	Del	Silent	Mut
> CL_001_pBEFO_CL_ssMRAH_MRAH-CH3.G1T4.v14hk	Match	yes	0	0	0	0
> CL_001_pBEFO_CL_ssMRAH_MRAH-CH3.G1T4.v14hk	Match	yes	0	0	0	0

従来: 朝(出社後)にデータコピーから手作業で実施していた

現在: 出社後すぐに次のステップの業務を開始できる

460時間/年の省力化を達成。夜間に解析が終了し、研究員は朝から実験結果に向き合うことができる。

ウェット研究員のデータ解析力を高める教育システムをドライ研究員が設計

研究員同士が教え伝えていくシステム



授業

受講者
研究員

次の講師に

授業

次の講師に

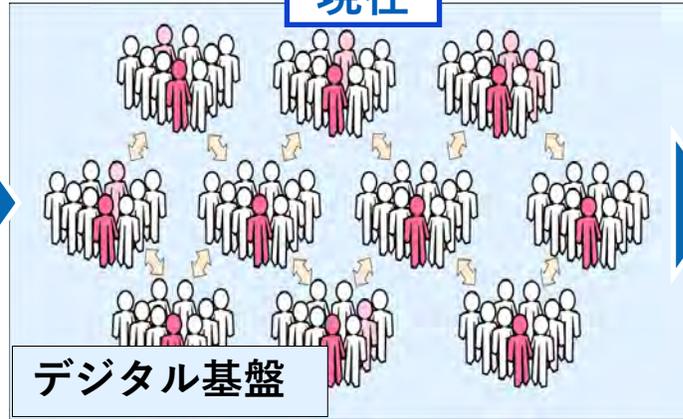
- ・ Pythonとデータ成型の基礎を学ぶ
- ・ ドリルと実問題解決のmix教材
- ・ 統合データ解析基盤ですぐ使えるスキルとなる

講師（データサイエンティスト）

DXに取り組む前

現在

研究所丸ごとDX



 : 一般研究員

 : 多少デジタルのわかる人財

 : デジタル人財

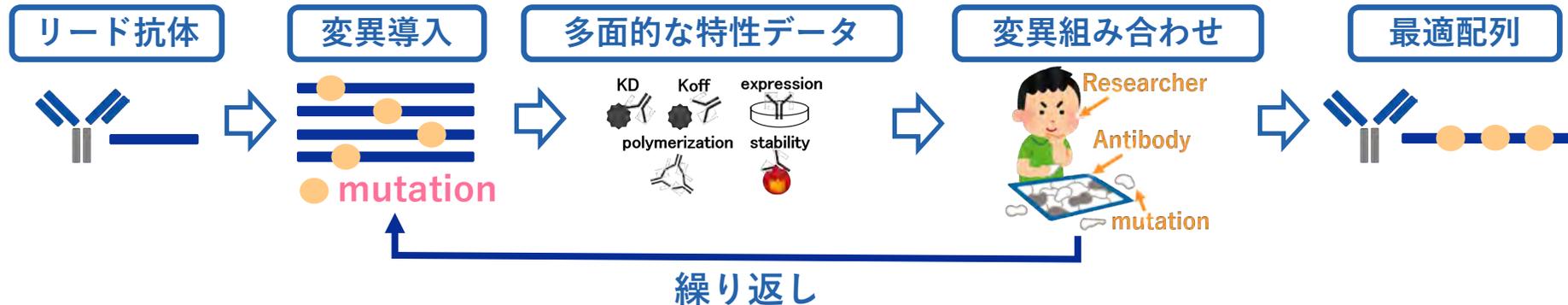
 : コミュニティ

AI創薬の実践：実験的労力のかかるプロセスにAIを用いる

抗体創薬プロセス



抗体最適化：リード抗体のアミノ酸の変異を繰り返し、最適な抗体を作製する

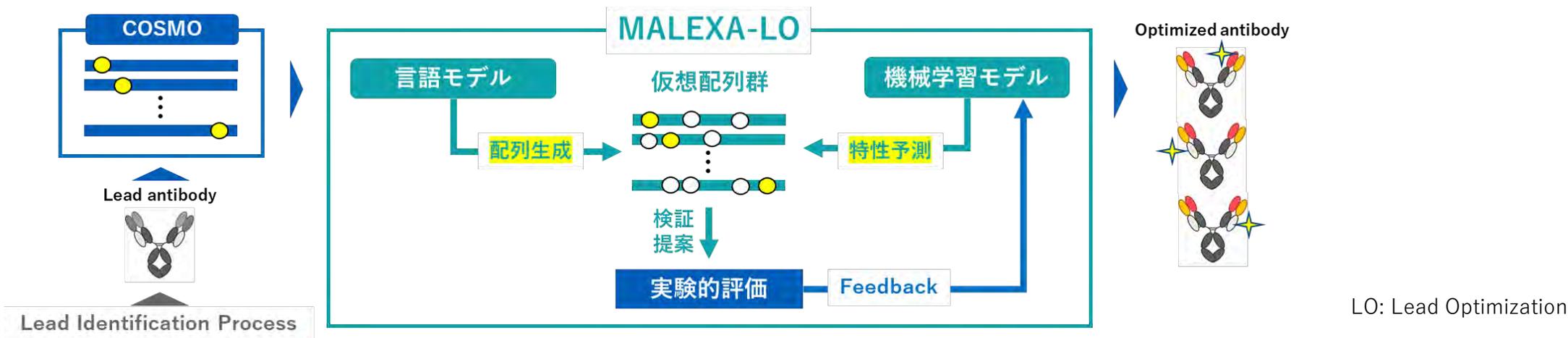


課題：変異の組み合わせは膨大である上、試行錯誤性が高い。→AIを使えないか

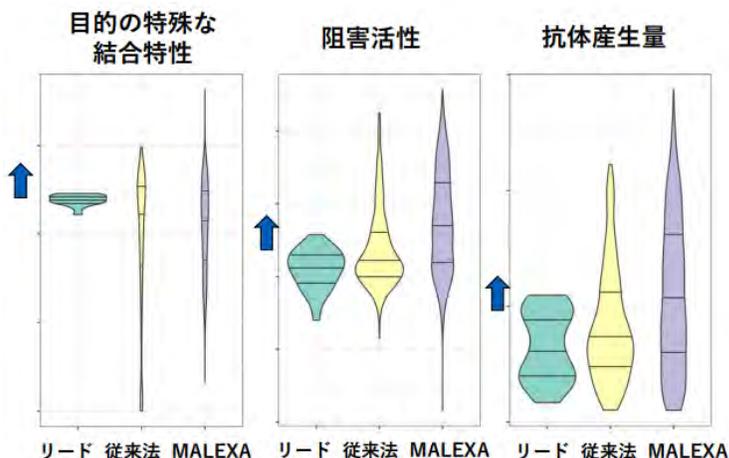
MALEXA[®]：機械学習を使った抗体配列デザイン

MALEXA[®] (マレキサ)：MAchine LEarning x Antibody

配列生成技術および特性予測技術を適用することで、機械学習による最適な抗体配列の導出を実現する

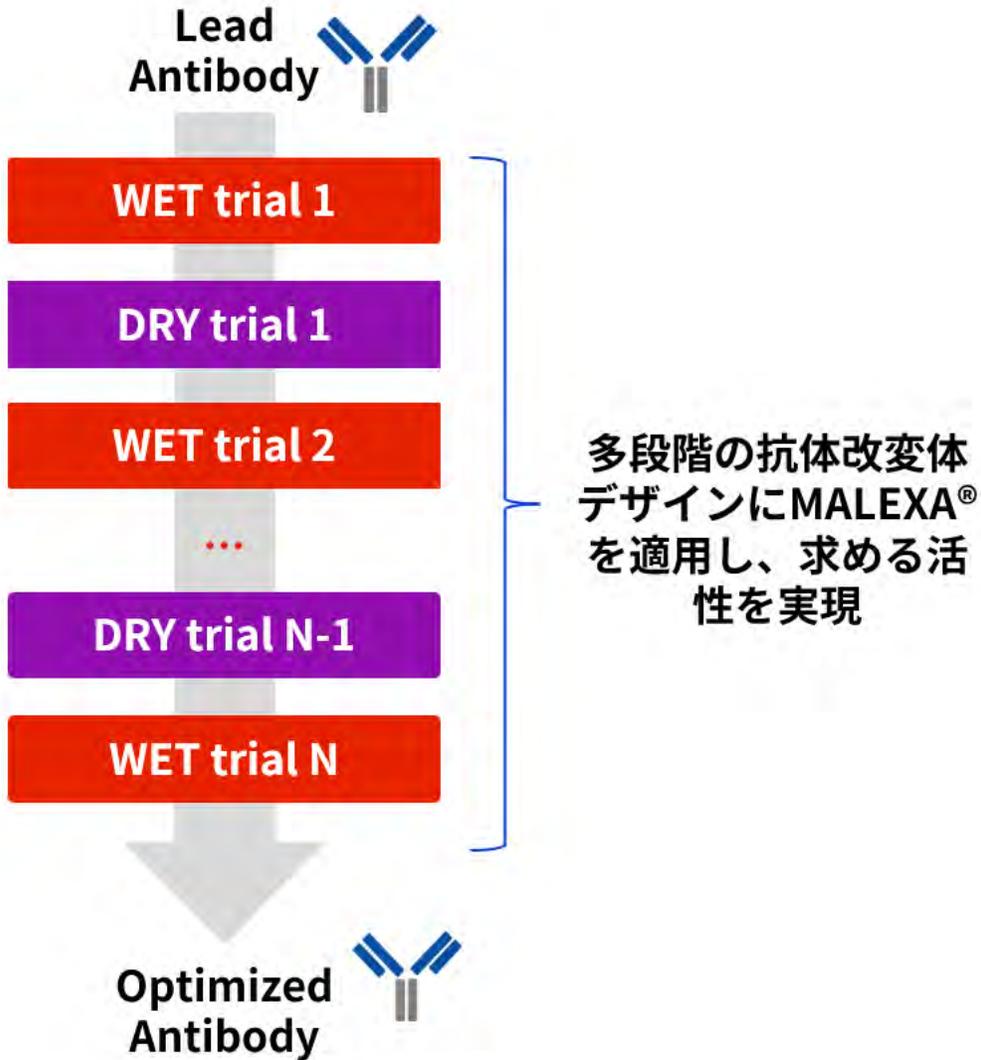


>各方法論で得られた抗体の*in vitro*における結合特性/阻害活性/抗体産生の分布

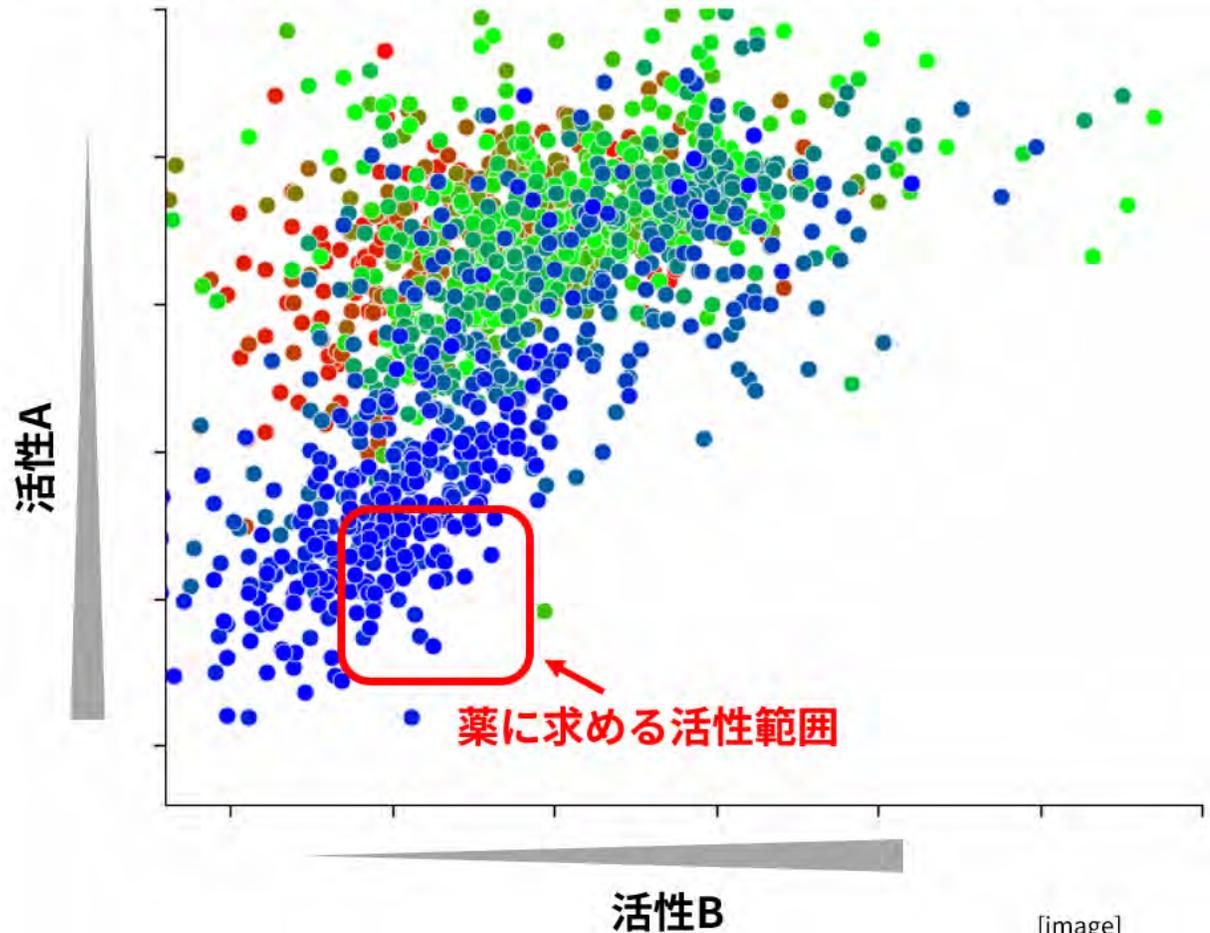


MALEXA[®]を用いることで、
従来法（＝研究員によるデザイン）よりも
優れた抗体配列を提案できることが確認された

AIによる複雑なデザインも可能になってきている



両立の難しい二つの活性を同時に最適化



DRYとWETの協調による実験プロセス構築にむけて

創薬リードタイム短縮 &新技術開発力強化

MALEXA®とLAシステムを連携した
Closed loopの実現

業務ボトルネックステップを
モバイル接続し抗体作製能力を倍増



まとめ(再掲)

- 弊社抗体創薬は、他社では模倣が難しいような抗体工学（技術ドリブン創薬）や品質に立脚した（Quality-centric創薬）、比較的モノづくり難易度の高い創薬を実践している。
- その中で、①Lab Automationによる実験自動化、②データ成型と蓄積、③研究員のデジタルリテラシー向上、④AIによる分子デザイン、の4つの項目でモノづくりに関わるDXに挑戦してきた。
- 弊社としてはDXには、単なる効率化を超えたヒトの思考を拡張する補完的な技術としての役割（質的高度化）を期待している。

創造で、想像を超える。





創薬DXPFエージェントによる AI駆動型創薬の実現

京都大学大学院医学研究科
理化学研究所計算科学研究センター
奥野 恭史





Deep Learning以降の情報科学・計算科学の進展

AIは創薬を変革できるのか

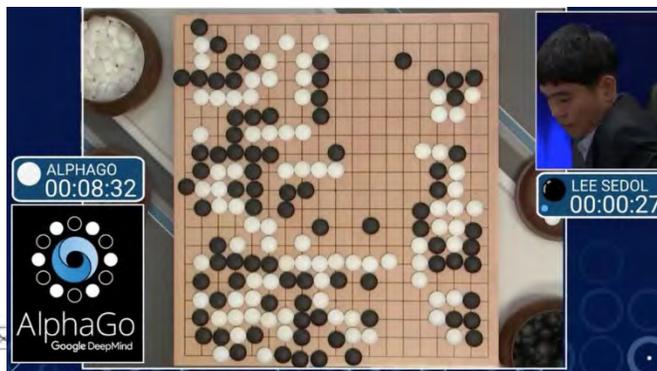
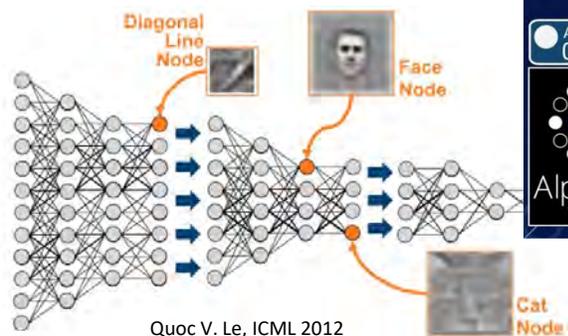
2022
ChatGPT

2025
AI Agent元年

2018
AlphaFold

2015
AlphaGo

2012
Deep Learning



<https://www.bbc.com/news/technology-35785875>



<https://alphafoldserver.com/welcome>



<https://yourstory.com/2023/05/everything-you-need-to-know-about-chatgpt>

最近のAI/計算機の躍進は半端ない

2012
スパコン「京」



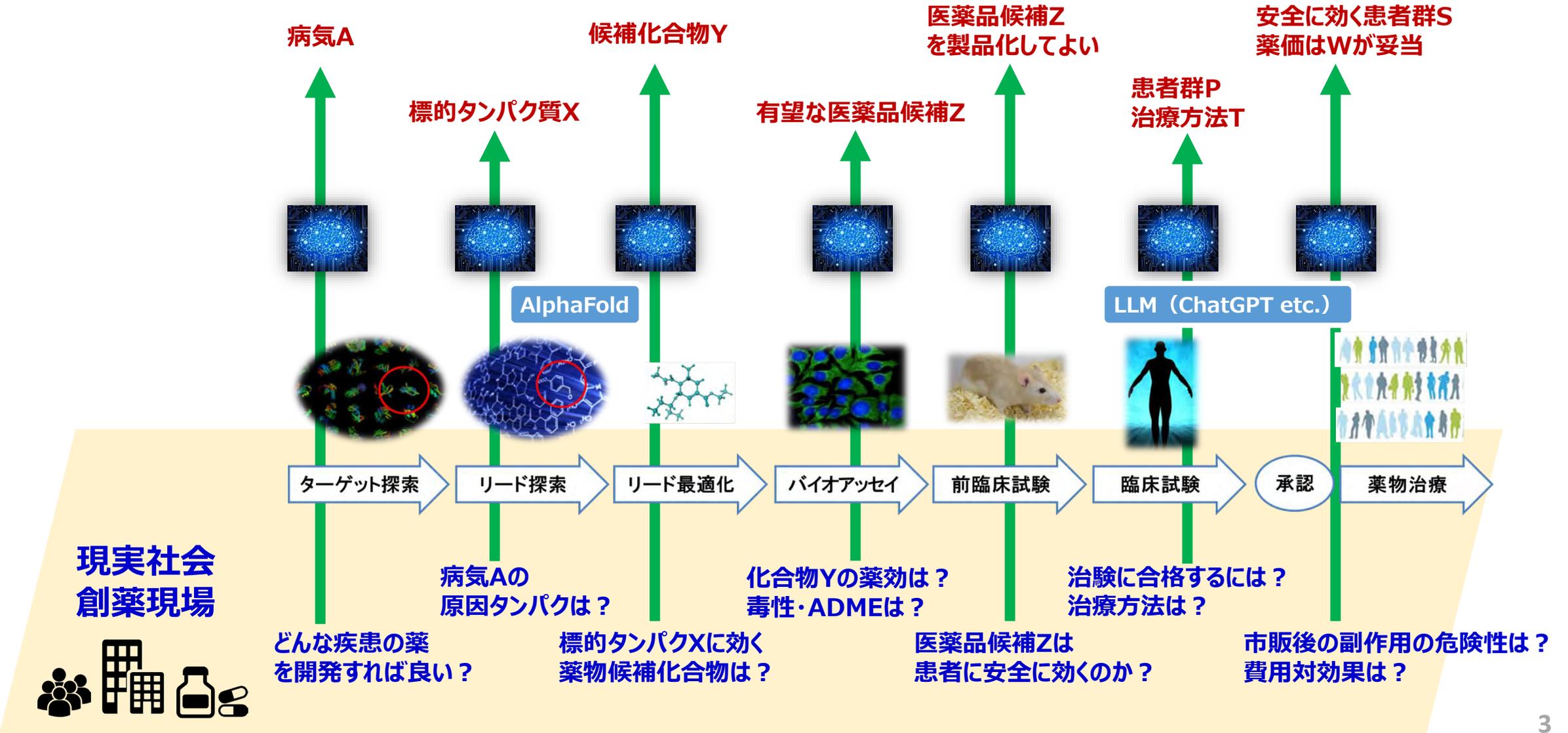
2021
スパコン「富岳」



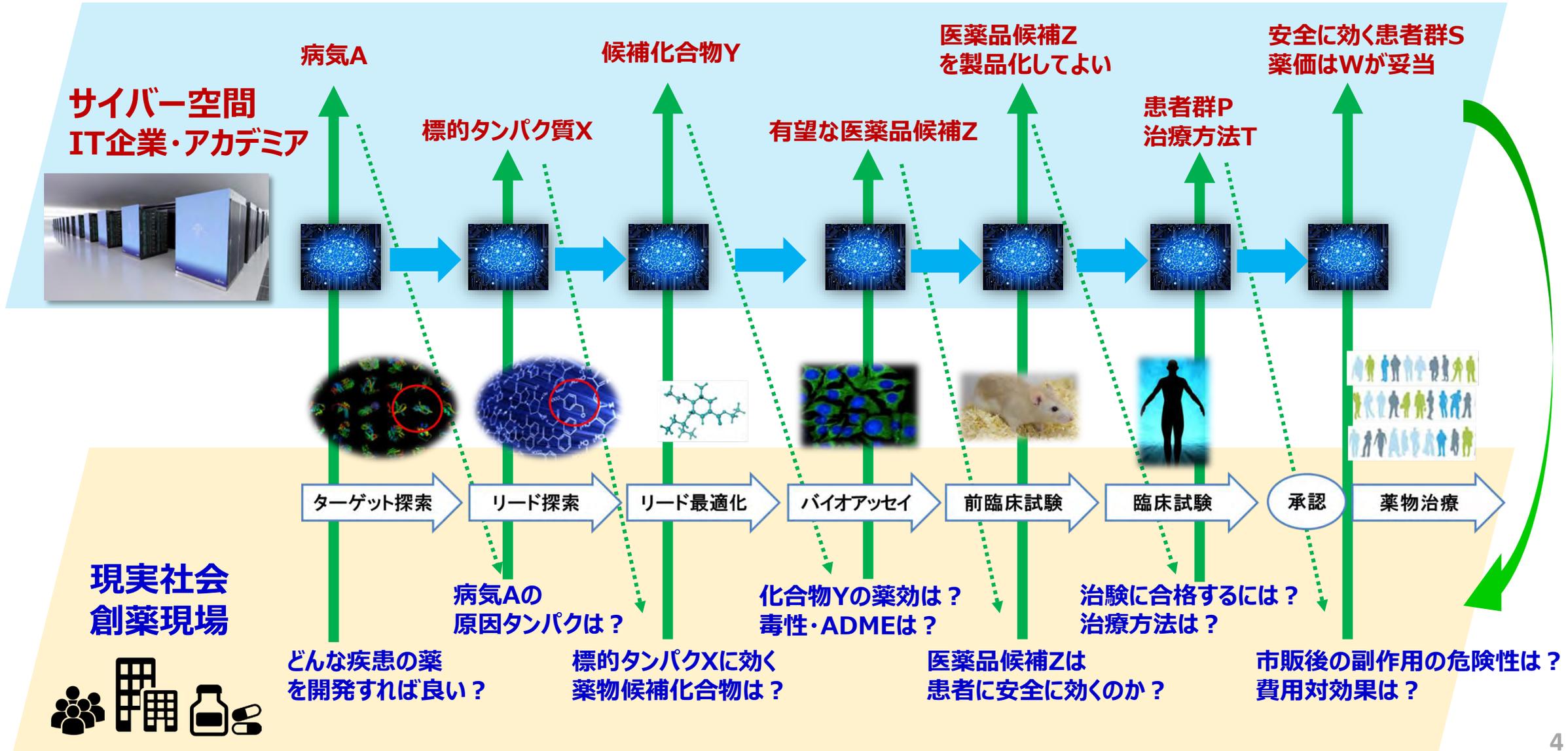
計算機の未来

- AIスパコン
- 富岳NEXT
- 量子コンピュータ
-

それぞれの部署や研究者が個々のアプリを使うだけでは創薬プロセスの高速化・効率化にはつながらない



創薬DXプラットフォーム

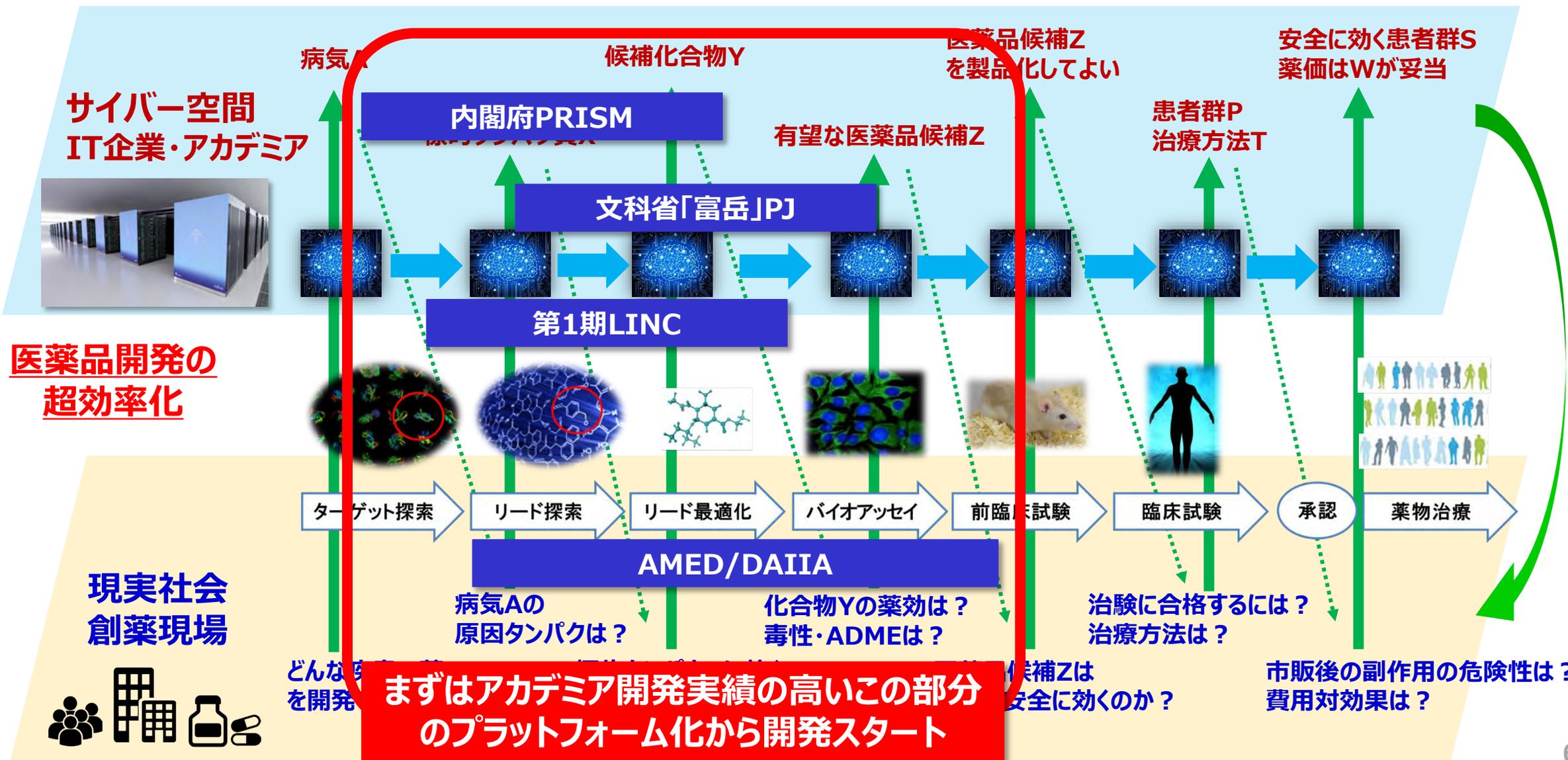


国民が医薬品へのアクセスを確保する、そして安心して使用できる環境とする、このためにこの我が国の創薬力向上することが不可欠ですが、その中で、昨今の創薬を取り巻く環境については、委員の方からも御指摘がありました。医薬品の開発の期間や費用が増大している、また確度の高い標的を発見することが困難になっている、こういった課題があると考えており、
御指摘の創薬DXプラットフォームのようなAIを始めとするデジタル技術は、こうした課題を解決し、我が国の創薬力向上につながる可能性を有している、このように考えております。



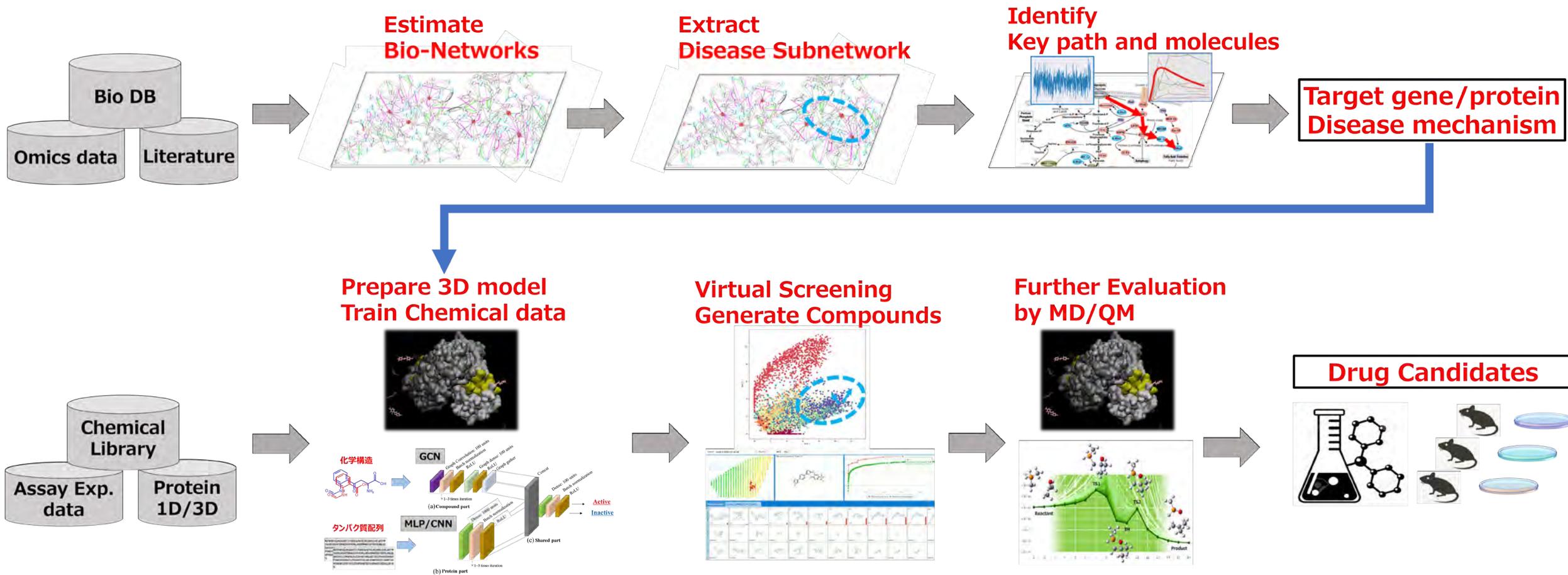
【国会中継】参院決算委 岸田首相・全閣僚出席で質疑（2024年4月1日）

<https://www.youtube.com/live/Y8JRg735rG0?si=QsUse8ayxdpT3893&t=19939>



- 「創薬ターゲット探索」⇒「リード化合物創出」に至る55種類のアプリからなる半自動計算フローを構築
- これまでに国プロ等が開発してきた技術を同じPFに、実装・連結・制御することで開発費、期間の劇的削減に成功

創薬ターゲット探索： 疾患名・患者サンプルデータ等を入力して、疾患メカニズムや標的タンパクを推定するHPC/AIフロー

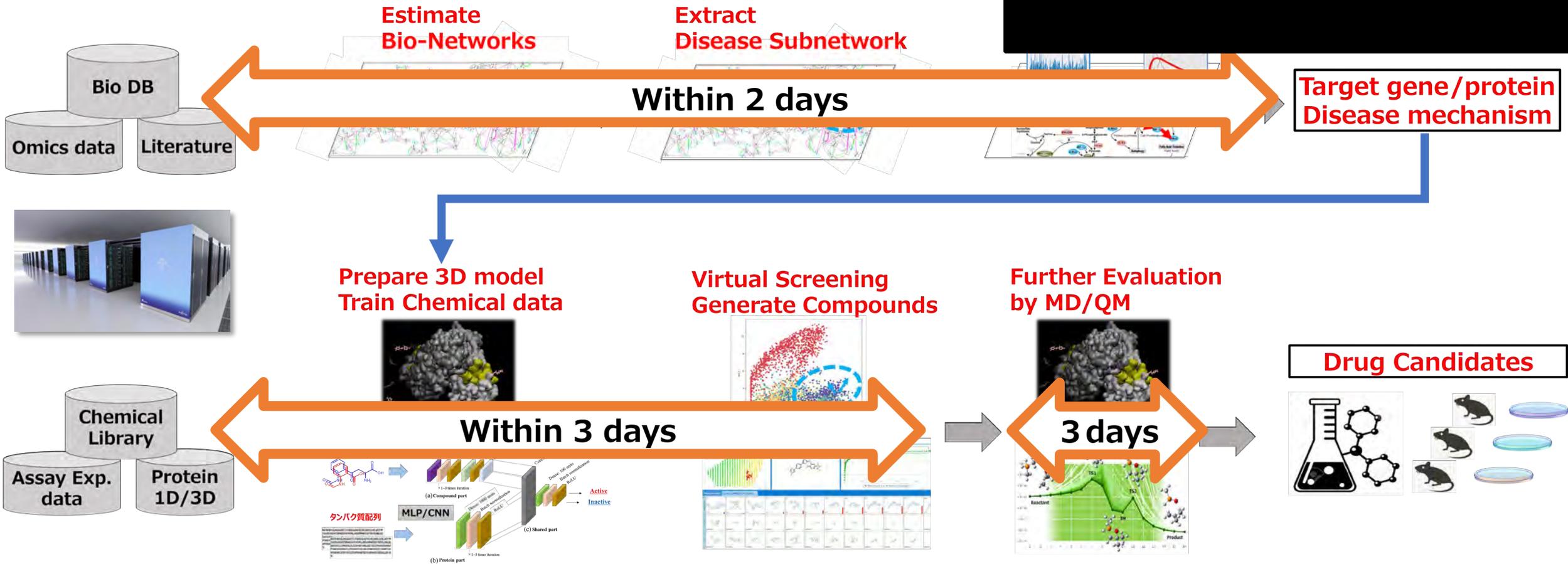


リード化合物創出： 標的タンパク質名を入力して、リード化合物を推定するHPC/AIフロー

創薬DXプラットフォーム

非常に高度な独自創薬計算アプリを含む55種のアプリを組み合わせて世界最高性能の予測精度と汎用性を実現

創薬ターゲット探索： 疾患名・患者サンプルデータ等を入力して、疾患メカニズムや標的タンパク質を探索

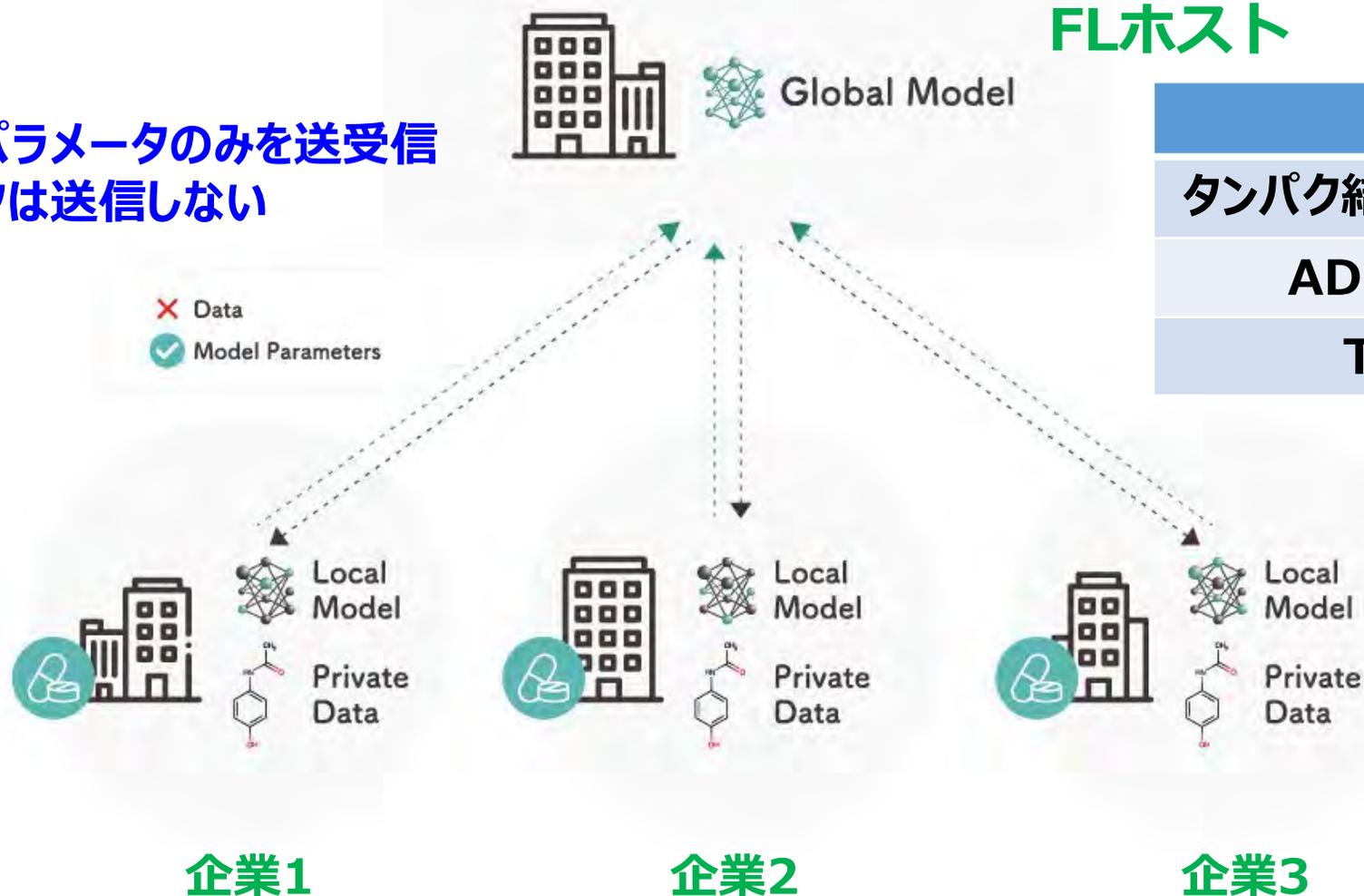


リード化合物創出： 標的タンパク質名を入力して、リード化合物を推定するHPC/AIフロー

AMED/DAIIA (2020~2024) での実績：製薬企業連合学習 企業データを秘匿しながら協調して高性能AIを構築

学習データを外部に持ち出さず、組織内に置いたまま、AIモデルパラメータのみを交換することで、複数組織が協調してAIモデルを学習することが可能

- AIモデルパラメータのみを送受信
- 学習データは送信しない



	公共DB	製薬企業
タンパク結合	1863万	1439万
ADME	5.8万	49万
Tox	1.7万	0.67万

事業期間中17社で実施
終了後現在10社が継続

<https://3dxfp-pub.med.kyoto-u.ac.jp/>

創薬DX Lite 次世代創薬AIデモシステム
Drug Discovery DX Lite

TOP 創薬DX Liteとは 注意事項 マニュアル 研究開発情報 お問い合わせ

創薬AIで予測する

症状から病名・疾患関連分子・薬候補を一貫して予測



創薬AIで予測する

症状から病名・疾患関連分子・薬候補を一貫して予測

症状・病名を入力してください

予測する

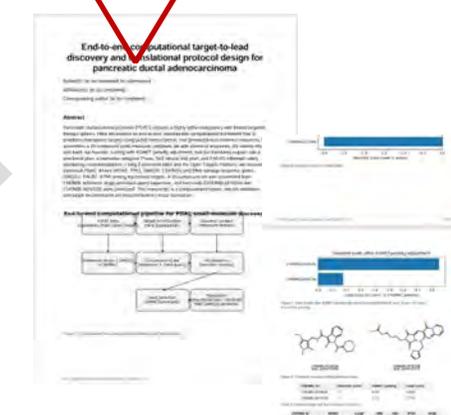


AIエージェントが、ユーザー入力に応じて、
ツール選択、計算フロー構築

AIエージェントが、適宜条件
検討しながら、計算を実行

AIが結果の解釈をし
レポート・論文作成

1180種以上の
アプリ



55種類の
アプリ



研究者が
創薬DXPF Webツールを開発

ユーザーが
創薬DXPF Webツールを操作

創薬DXPFエージェントと AI for Science



AIによる
仮説生成



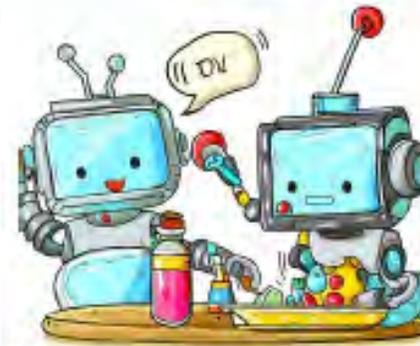
AIによる
計算・実験プロトコル
作成

AIによる
論文学習



創薬DXPF

実験ロボット



AIによる
解析結果の
解釈・論文化



AIによる
計算・実験結果の解析

入力文：
マルチオミクスデータ、臨床データを利用して、すい臓がんの創薬標的分子を探索、同定し、分子生成した化合物20個からヒット化合物を探索、ADMET予測に基づく最適化を行い、前臨床評価、成功確率の高い治験プロトコルを作成、さらに実臨床での有害事象リスク推定し、最終的に実臨床での治療プロトコルを決定したい

Example Execution Flow

The screenshot shows the '3dxf Designer Agent' interface. On the left, there is a 'Run' button and a 'Mode' section with '非対話 (Batch)' and '対話 (Interactive)' options. Below that is a 'Prompt' section containing the input text. At the bottom left, there are 'LLM設定' and '実行オプション' sections. The main area is titled 'Workflow' and shows a search bar for logs. A large black area covers the log output, with the text 'No stdout logs yet.' visible.

Manuscript-Style Report Generated from the Results

End-to-end computational target-to-lead discovery and translational protocol design for pancreatic ductal adenocarcinoma

Author(s): [to be completed for submission]
Affiliation(s): [to be completed]
Corresponding author: [to be completed]

Abstract

Pancreatic ductal adenocarcinoma (PDAC) remains a highly lethal malignancy with limited targeted therapy options. Here we present an end-to-end, reproducible computational framework that (i) prioritizes therapeutic targets using public transcriptomic and genetic/clinical evidence resources, (ii) assembles a 20-compound small-molecule candidate set with chemical structures, (iii) selects hits and leads via heuristic scoring with ADMET-penalty adjustment, and (iv) translates outputs into a preclinical plan, a biomarker-adaptive Phase 1b/2 clinical trial draft, and FAERS-informed safety monitoring recommendations. Using Expression Atlas and the Open Targets Platform, we recovered canonical PDAC drivers (KRAS, TP53, SMAD4, CDKN2A) and DNA damage response genes (BRCA2, PALB2, ATM) among top-ranked targets. A 20-compound set was assembled from ChEMBL reference drugs and class-query expansion, and two leads (ChEMBL5918534 and ChEMBL5876728) were prioritized. This manuscript is a computational report: wet-lab validation and target deconvolution are required before clinical translation.

End-to-end computational pipeline for PDAC small-molecule discovery

```
graph TD
    A[Public data  
(Expression Atlas, Open Targets)] --> B[Target prioritization  
(rank aggregation)]
    C[Genomic context  
(literature fallback)] --> B
    B --> D[Reference drugs + SMILES  
(ChEMBL)]
    B --> E[20-compound set  
(reference + class query)]
    D --> E
    E --> F[Hit selection  
(heuristic scoring)]
    F --> G[Lead selection  
(ADMET-penalized)]
    F --> H[Translation  
(preclinical plan, trial draft,  
RWE safety & protocol)]
```

Figure 1 | Overview of the computational workflow and deliverables.

生物領域

胃がんネットワーク標的探索

```

Problems Output Debug Console Terminal Ports
print("\nUpdated plan:")
plan = [
    "1. [x] Tool availability check",
    "2. [x] advanced_literature_search_agent failed",
    "3. [x] Dataset selected and downloaded (GSE84437)",
    "4. [x] Parsed expression + metadata",
    "5. [x] Normalized expression for BN learning",
    "6. [x] Reduced dimensionality to top variable genes",
    "7. [x] Learned BN structure (single fit)",
    "8. [x] Bootstrapped edge confidence; built consensus network",
    "9. [x] Identified master genes; ranked top K",
    "10. [x] Probe->gene symbol mapping (GPL6947)",
    "11. [x] UniProt mapping",
    "12. [x] OpenTargets overlap annotation",
    "13. [x] HPA stomach expression (top10; optional)",
    "14. [x] Final outputs saved"
]
print("\n".join(plan))

print("\nKey final files in AGENT_OUTPUT_DIR:")
for p in sorted(agent_dir.glob("*")):
    if p.is_file():
        print("-", str(p))
[agent] exec_end
[agent] observe_begin
<observe>
[execution] language=python ok=False returncode=1
----- STDOUT -----
Proceeding with Step 12-14: UniProt mapping, OpenTargets overlap, and optional HPA stomach expression.

Gene symbols available for mapping: 22
Symbols: ['CNN1', 'MYH11', 'DES', 'C2orf40', 'PGA3', 'PGA5', 'CDH17', 'CDX1', 'ACTG2', 'GAGE12I', 'GAGE2B', 'GAGES', 'GAGE12G', 'TFF1', 'CLDN18', 'CA2', 'TFF2', 'MGP', 'CTSE', 'TAGLN']

```

- 胃がんRNAデータ取得
- ベイジアンネットワーク解析
- マスター遺伝子特定
- 標的マッピング

化学・物理領域

新型コロナウイルス化合物探索

```

Problems Output Debug Console Terminal Ports
## Candidate solution text (raw; resume 実行では差分の可能性あり)
## 最終レポート: mPro ドッキング → 5 ps MD → ADMET 予測 → 特許性 (PubChem 化合物 5 件)

### ブランチェックリスト (完了)
1. [x] Tool Finder_LLM を使用してツールチェーンを確認 (完了)
2. [x] query_pubchem により PubChem から候補分子 5 件を取得 (完了)
3. [x] ちょうど 5 CID に解決/重複排除 (完了)
4. [x] mPro 受容体 (6LU7) をドッキング用に準備 (完了; 必要な修正を適用)
5. [x] リガンドを準備 (SMILES+3D+PDBQT) (完了)
6. [x] 共結晶のヘテロ原子からドッキングボックスを定義 (完了)
7. [x] 5 リガンドすべてのドッキングを実行 (完了; 必要な修正を適用)
8. [x] ドッキングスコアで順位付け (完了)
9. [x] MD をセットアップ (制限ありで完了; 下記参照)
10. [x] 5 ps MD を実行 (制限ありで完了)
11. [x] MD を後処理 (制限ありで完了)
12. [x] ADMETAI 物性 + 毒性を実行し集約 (完了; パースを修正)
13. [x] PubChem 特許関連付けによる特許性評価 (完了; パースを修正)
14. [x] 任意の PubChem メタデータクロスチェック (同義語) (完了)
15. [x] 最終順位レポートを作成 (本レポート)

---

## 1) 選定化合物 (PubChem)
入力セット (CID, SMILES) は以下に保存:
- **/home/okuno/project/3dxfp_bm_tu_designer_260106_v2/3dxfp_runs/20260116_100732_after_performing_doc_king_calculations_with_mpro_fo_4462361f/candidate_01/step3_pubchem_5_dedup.csv**

CIDs:
- 6505527
- 110778773
- 99796981

```

- 化合物・タンパク構造取得
- ドッキング計算
- MD計算
- ADMET予測
- 特許性評価

臨床領域

AKI発症予測

```

bash
export AKI_OUT_DIR="/home/okuno/project/3dxfp_bm_tu_designer_260106_v2/3dxfp_runs/20260116_000704_obta_in_public_ehr_and_make_an_early-onset_predicti_fe2a73dd/candidate_02/final_outputs/repro_run"
python "/home/okuno/project/3dxfp_bm_tu_designer_260106_v2/3dxfp_runs/20260116_000704_obtain_public_ehr_and_make_an_early-onset_predicti_fe2a73dd/candidate_02/final_outputs/aki_early_onset_prediction_synt_hca.py"

### 最終チェックリスト (PLAN 2をここで可能だった範囲にマッピング)
1. [✓] 公開ダウンロード可能なEHR風データセットを特定 (Synthea合成EHRを使用; 候補リストを保存)
2. [✓] データセット+スキーマをダウンロード (CSVをダウンロードして抽出; スキーマはヘッダから)
3. [✓] データを読み込み/プロファイル; タイムスタンプ確認; AKIシグナル特定 (クレアチニンあり)
4. [✓] HPO妥当性確認 (ツール実行)
5. [✓] KDIGOベースのラベリング (クレアチニンのみ近似) + 時間的カットオフ
6. [✓] 前処理+不規則時系列からの特徴量抽出 (0-6h集計)
7. [✓] ベースライン+木系モデルを学習
8. [✓] CV+ブートストラップCIで評価 (コホート極小である点を明記)
9. [✓] 任意: FAERS腎毒性シグナル探索
10. [X] PubMedガイドラインツールが失敗 (api_keyが必要)
    [✓] 変更: NICE/MHOガイドライン検索を完了
11. [✓] 完全再現可能スクリプト+モデルカードを出力

## 実 (非合成) のAKI早期警告モデルに向けた推奨
臨床的に意味のあるモデルが必要であれば、次のステップはアクセス取得後に実ICU EHRデータセット (例: MIMIC-IVまたはeICU) 上で同じパイプラインを実行することです。理由は以下の通りです:
- KDIGOには尿量と透折指標が重要。
- ベースラインクレアチニン定義は慎重な取り扱いが必要、
- 安定した評価とキャリブレーションには十分なサンプルサイズが必要。

クレアチニン+尿量+受診タイムスタンプを含む実データの派生コホート (CSV/Parquet) をアップロードいただければ、ラベリングを完全なKDIGOに適合させ、堅牢な時間ベースCVとキャリブレーションプロットを作成できます。

```

- EHRダミーデータ取得
- アウトカム定義・ラベル付け
- 特徴選択
- 機械学習モデル構築
- 交差検証
- FAERS参考情報調査

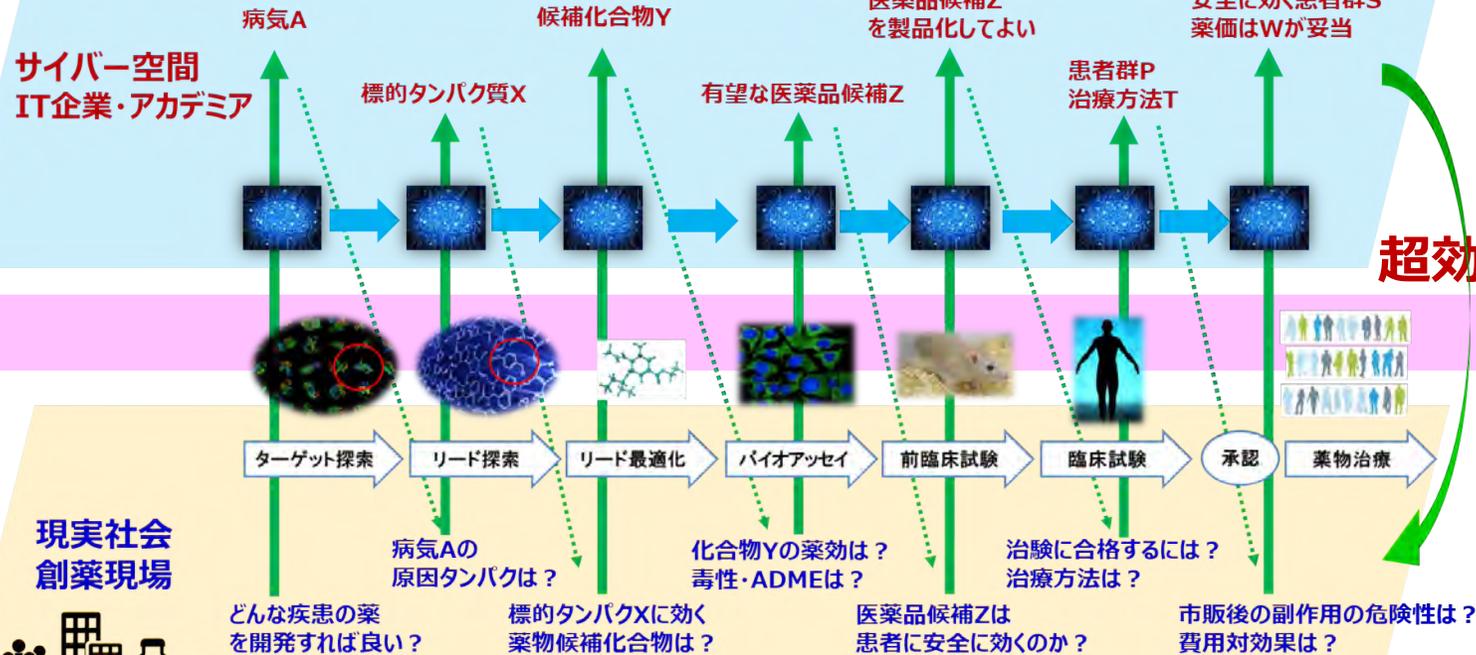
すべて1時間以内に完了

2025年度からAMED新規事業にて
抗体・核酸・ペプチド系創薬AIへ拡張

前臨床・臨床へのさらなる拡張：
予算措置が必要

AI/DX人材育成
PFを用いたOJT

製薬企業
ベンチャー
アカデミア



超効率的・高速・低コストの
画期的新薬



AI for Science : 創薬・医療から生命科学全体へ

他機関・施設連携

データベース
NBDC・DBCLS
京大病院
国際医療研究センター
医薬基盤研

化合物ライブラリー
東大創薬機構
京大薬

大型実験施設
理研Spring8/SACLA
阪大蛋白研 cryoEM

ゲノム・オミクス解析
東大医科研
京大薬

バイオバンク
京大病院バイオバンク
東北MMB



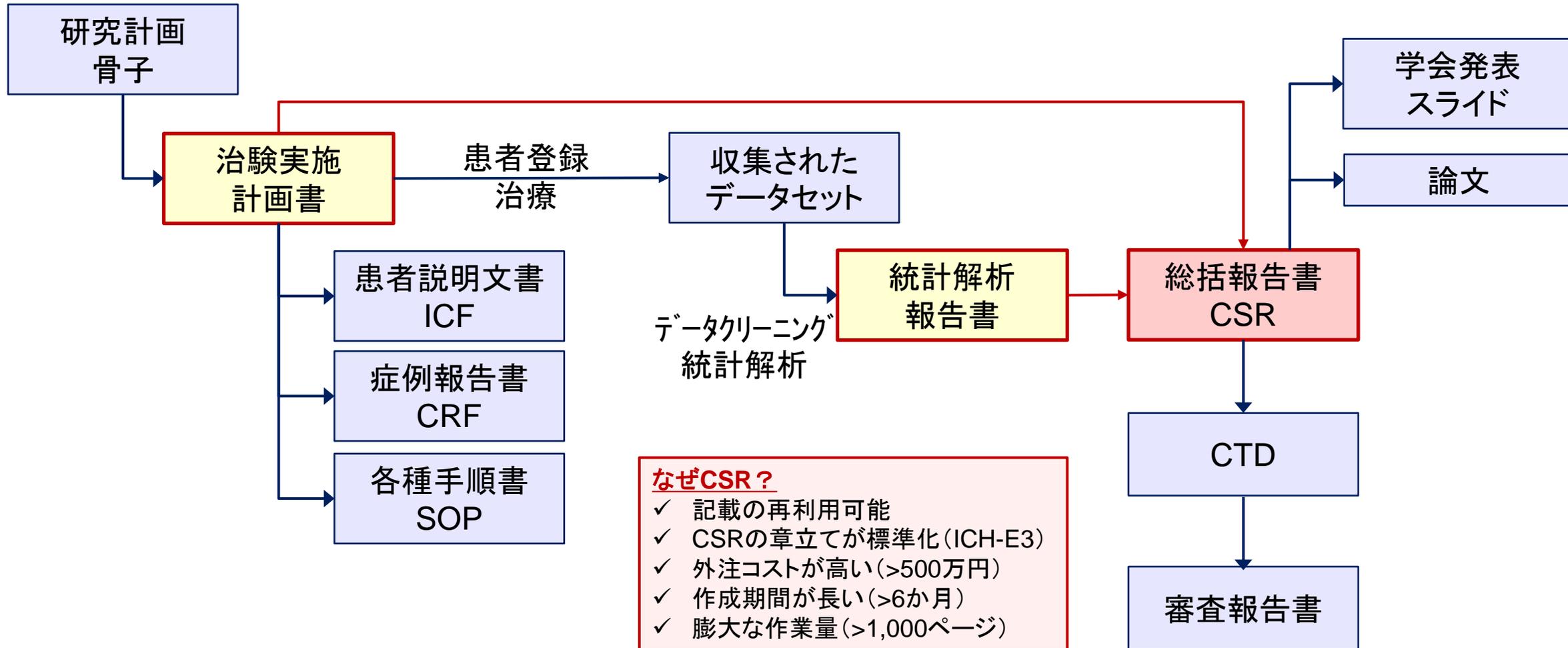
臨床試験プロセスにおける 生成AIの利活用

国立がん研究センター中央病院
国際開発部門/臨床研究支援部門
中村 健一

医薬品開発におけるAI/DX



臨床試験プロセスは文書のつながり



治験関連文書の自動作成プロジェクト

■ 第1期(2024年度)

- 治験総括報告書(CSR)の自動作成プロジェクトを開始
- ベンダーとの共同研究で、生成AIを用いたCSR作成ツールを開発
- CSRの第9章についてAIによる自動作成の精度を評価

■ 第2期(2025年度)

- 第1期の結果を踏まえ、CSR作成用プロンプトの改善と汎用化を実施
 - CSRの第1章、第5～8章、第9章について自動作成
- PMDAと共同で厚労科研を獲得(2025年11月～)
 - 「生成AIを活用した薬事承認申請・審査関連文書作成の推進のための研究」
 - NCCではCSR作成のさらなる精度向上が主目的
 - PMDAでは審査報告書の一部の自動作成に着手
- 並行してNCCではプロトコールの自動作成にも着手

文書自動作成ツールの仕組み

例：CSR



過去の文書セットをRAGで参照させ文書を生成

*RAG: Retrieval-Augmented Generation

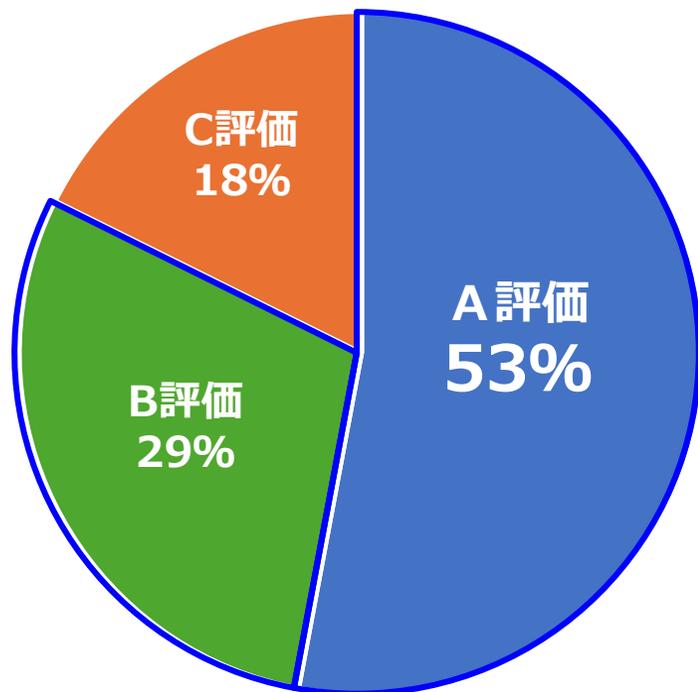
実際のCSR自動作成のイメージ

The screenshot displays the 'うくろAI' web application interface. On the left, a sidebar shows 'Reference Link' and 'Generated Options' (Option 1 and Option 2). The main area is an 'エディタ画面' (Editor) showing a CSR document titled '9. 治験の計画'. A red dashed arrow points from the 'Reference Link' to a 'Data source' box on the right. A red circle highlights the 'Generated Options' section, with a callout stating '自動作成された複数の文書候補が表示されるため、好みに応じて選択可能'. A blue box highlights the 'エディタ画面' area, with a callout stating 'エディタ画面で追記・修正が可能 (赤枠：追記した箇所)'. A green box highlights the 'チャットボット' (Chatbot) area. A yellow highlight in the document is labeled '参照箇所を黄色ハイライトで表示' (Display reference location with yellow highlight). The document content includes sections like '9.1 治験の全般的デザインおよび計画-記述' and '9.1.1 試験デザイン'.

自動生成されたCSR：赤枠部分に複数の文書候補が生成され、好みの候補を選択可能。
内容は青枠のエディタ画面で自由に編集することが可能。
緑枠のチャットボットで一般的な質問などが可能。

自動作成したCSRの精度 (第9章)

第1期 (評価項目数119)



Phase1 (試験固有)
A + B = 約80%が実用レベル
 (ほぼそのまま使用可能)

Phase2 (汎用化)
A + A- + B = 約97%が実用レベル
 (修正最小限で使用可能)

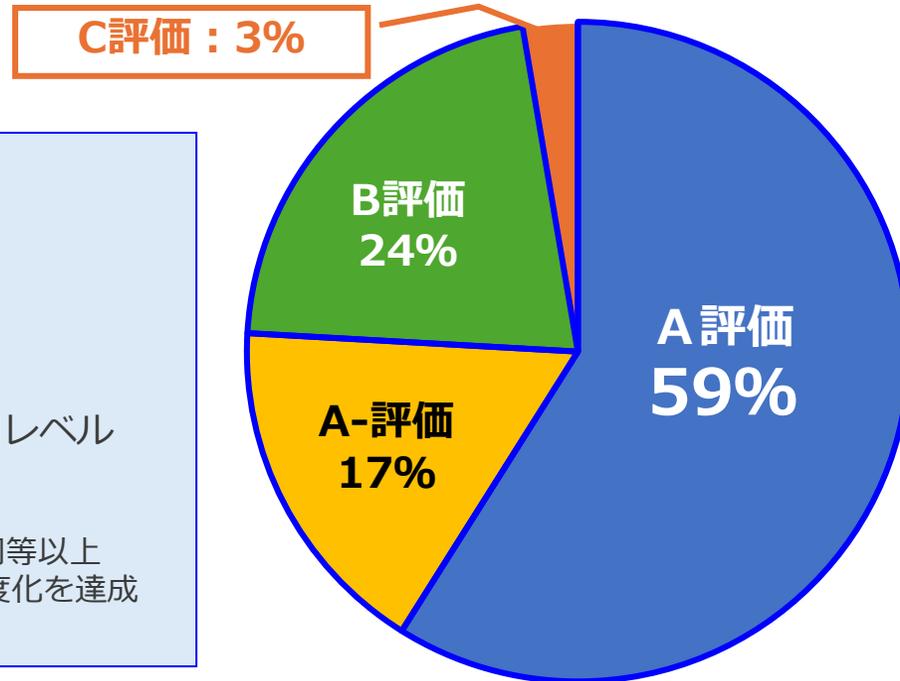
➡汎用化後も生成精度はPhase 1と同等以上
 実用性を損なうことなくツールの高度化を達成

評価項目	評価内容	項目数	割合
A	ほぼそのまま利用可能	63	53%
B	微修正で利用可能	35	29%
C	大幅な追加作業が必要	21	18%
D	利用は困難	0	0%

【C評価の例】

- 表の形式が崩れる
- 検査の評価期間のハルシネーション
- 実施していない「薬物濃度の測定」の記載を生成

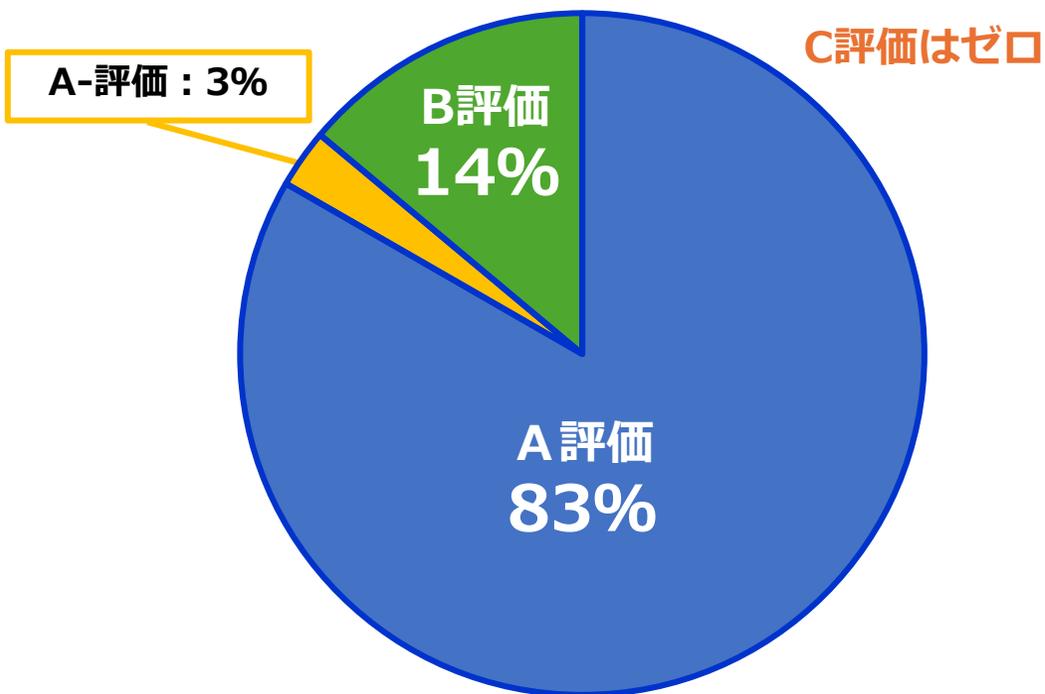
第2期 (評価項目数112)



評価項目	評価内容	項目数	割合
A	そのまま利用可能	66	59%
A-	軽微な記載整備で利用可能	19	17%
B	微修正で利用可能	24	24%
C	大幅な追加作業が必要	3	3%
D	利用は困難	0	0%

自動作成したCSRの精度 (第1章・第5～8章)

第2期 (評価項目数36)



■ 第2期の結果

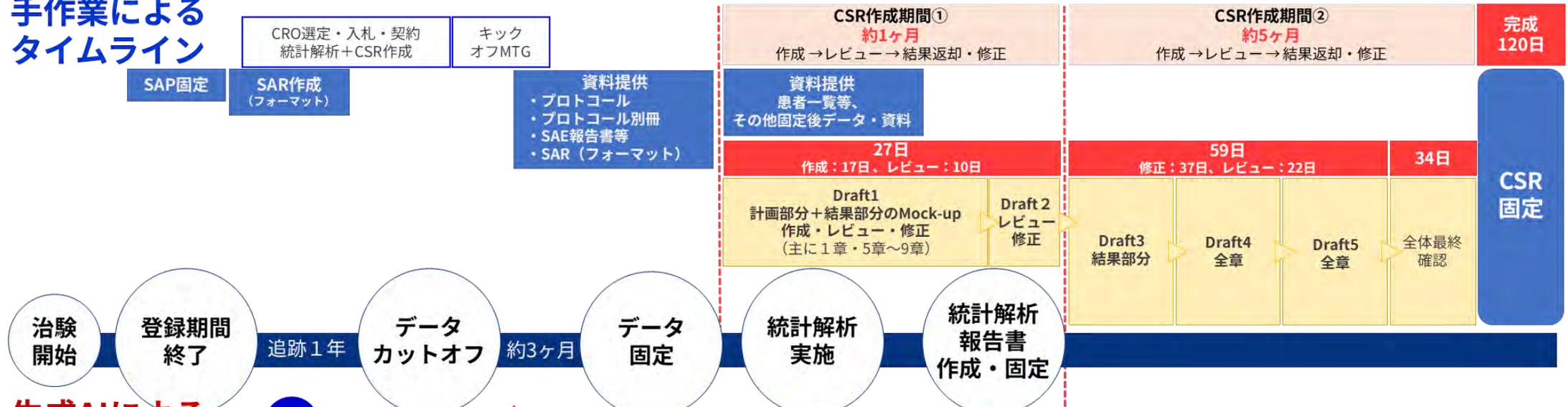
- 第1章・第5～8章は**100%**、第9章でも**97%**が実用可能と評価され、最小限の微修正で利用可能であることを確認
- 今後、複数の試験での再現性を評価しつつ、がん以外の臨床試験での精度を評価予定
- インプット側のプロトコール等の章立ての標準化を進めることも精度向上のためには必要

評価項目	評価内容	項目数	割合
A	そのまま利用可能	30	83%
A-	軽微な記載整備で利用可能	1	3%
B	微修正で利用可能	5	14%
C	大幅な追加作業が必要	0	0%
D	利用は困難	0	0%

作業時間の短縮効果

生成AIの活用により営業日換算で120日→78日に短縮

手作業による タイムライン



生成AIによる 自動作成



キックオフを早い段階で実施でき、データカットオフ時点で提供可能な資料 (Draft含む) を提供

約1.5カ月でプロンプトチューニングが完了するため、データ固定前にCSRドラフトの作成開始が可能

生成AIの利活用事例：患者-試験マッチング “Patient to Trial” type

■ 個別患者に適合する臨床試験の自動マッチング

□ TrialGPT(患者-試験マッチングLLM)

■ Retrieval: 大規模フィルタリングによる候補試験の抽出

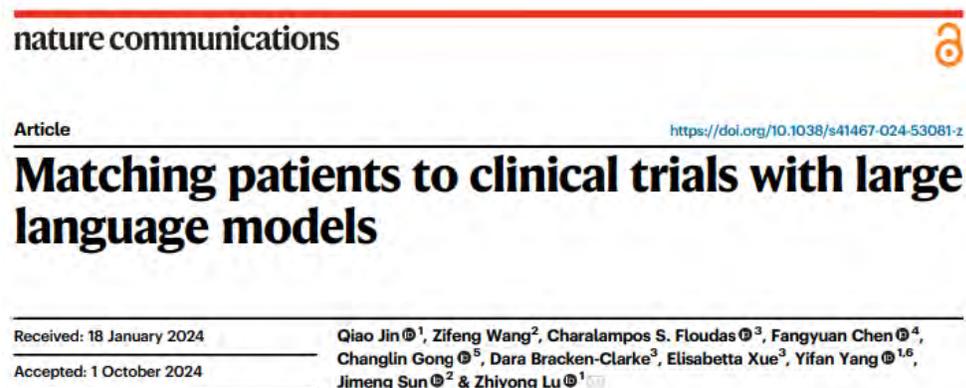
■ Matching: 電子カルテ内の患者情報と適格規準との関連性を評価

■ Ranking: 試験レベルで関連性をスコア化し、総合スコアで試験をランキング

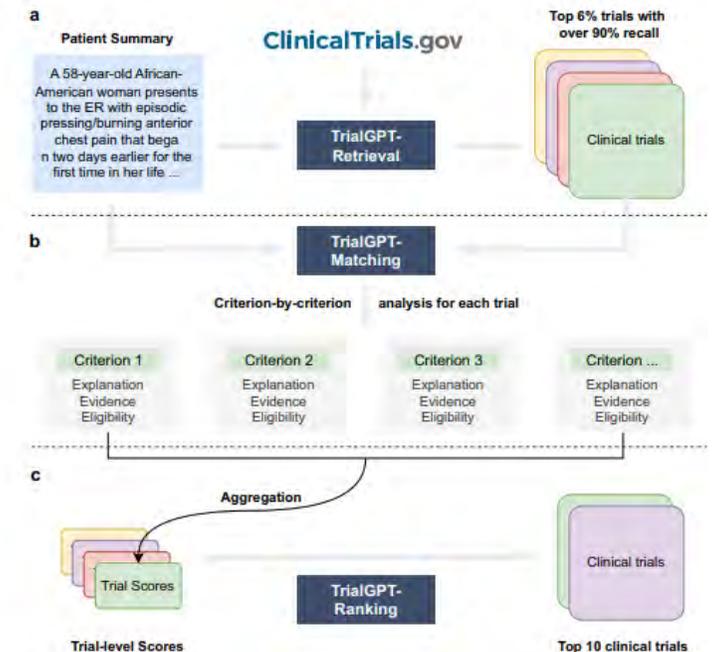
□ 患者-試験マッチング判定の精度は87.3%(専門家の水準に近いレベル)

□ 試験ごとのランク付けの精度は43.8%(従来モデルを上回る)

□ 患者1人あたりの試験照合に要する時間が42.6%短縮



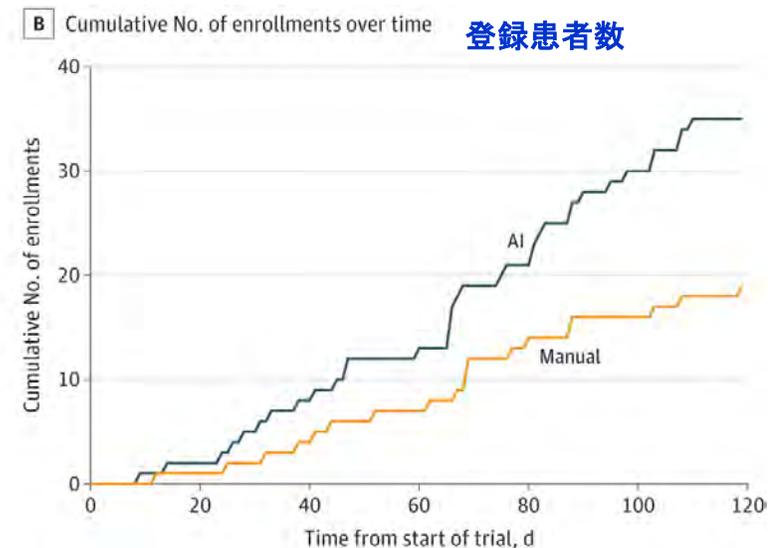
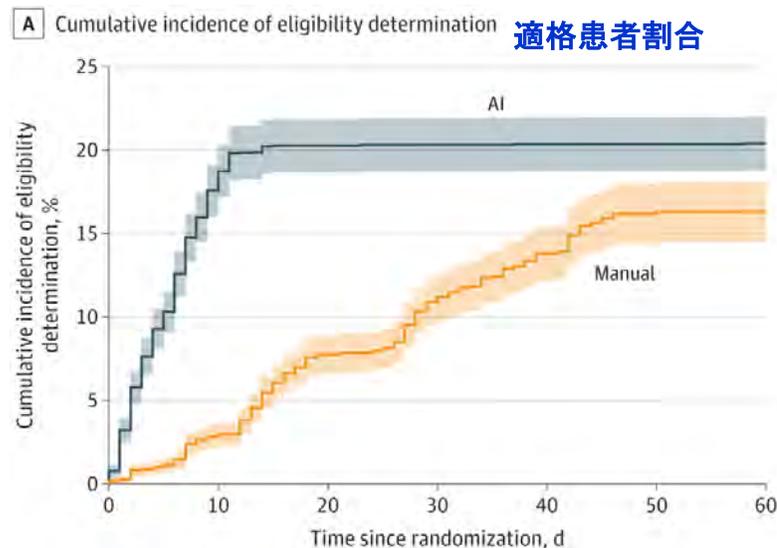
Jin Q, et al. Nat Commun 15, 9074. 2024



生成AIの利活用事例：患者-試験マッチング “Trial to Patient” type

■ 特定試験における適格患者のスクリーニング

- 電子カルテ上の非構造化診療メモから、臨床試験の適格規準に合致する患者を自動抽出
- 心不全治療の臨床試験(COPILOT-HF試験)において、AIツール使用群と従来の手作業群にランダム化して適格性判定までの時間を比較
 - AI群で適格性判定を有意に迅速化(ハザード比1.78, $p < 0.001$)
 - 適格と判定された患者の割合: AI群 20.4% > 手作業群 12.7% ($p < 0.001$)
 - 実際の試験への登録患者数: AI群 35例(1.6%) > 手作業群 19件(0.9%) ($p = 0.04$)



生成AIの利活用事例：予測系

■ 臨床試験成功確率の予測

- GPT-4を用いた臨床試験結果予測モデル
 - 430件の過去試験データと過去論文からモデルをファインチューニング
 - 強化学習で専門家のフィードバックを組み込み
- 試験成功/不成功を70%以上の精度で的中 (AUC=0.84)
 - 人間では気付かない試験設計上の欠陥を指摘したケースも

<https://www.panda-int.com/insights/ai-in-clinical-trials/>

■ 施設選定の最適化

- 多発性骨髄腫の試験を題材に、全米1,246施設・約15,881人の患者データをベースに、目標登録ペースと患者の多様性の規準に合致した施設選定案を提示
- 3施設のみを選定することで、目標登録ペース(10例/月)を上回りつつ、患者の多様性を確保できることを予測
- 従来の手作業よりも迅速・高精度に施設選定が可能であることを示唆

Ruchlin I, et al. J Clin Oncol 42, 2024 (suppl 16; abstr 1558)

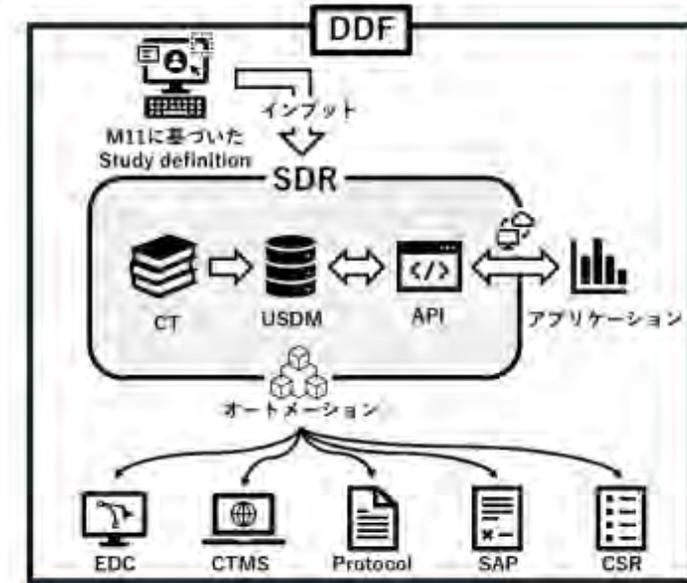
ICH-M11とDigital Data Flow

■ ICH-M11による標準化

- なぜプロトコルの国際標準化？
 - 企業別、団体別に異なるプロトコルテンプレート
 - プロトコル情報を下流のCRFやCSRへ転記する際のエラー
 - 手作業によるプロトコル作成で生じる遅延
- これらを解決するためプロトコルを国際標準化 (CeSHaRP)
 - プロトコル情報のデータ互換を可能とする技術仕様を定義
 - 2025年11月にICH-M11としてStep 4に到達

■ Digital Data Flow (DDF)

- TransCelerate主導の、臨床試験情報の構造化・標準化プロジェクト
- Study Definitions Repository (SDR)を最上流とする、プロトコルやEDC、CTMS、あらゆる下流の文書への自動的な流し込み
- 構造化されたデータはAPIを通じて外部連携も可能に
- DDFと生成AIの組み合わせにより、試験期間の飛躍的な短縮とコスト削減が期待される



製薬協医薬品評価委員会データサイエンス部会
「ICH M11導入後の医薬品開発への影響」

✓ 生成AI利活用を推進する上でも、プロトコルの標準化と臨床試験情報の構造化は極めて重要

最後に

- 医薬品開発のあらゆるプロセスでAI利活用が進む
 - 治験関連文書の自動化は最もハードルが低い利活用事例
 - CSRをはじめとした文書自動作成により作業時間短縮、コスト削減が進む
 - 文書作成のみならずデータマネジメント、統計解析等、あらゆるプロセスにAIの適用が進む
 - 米国、中国等ではプロセスの「一部」ではなく「全体」のautomationがなされつつある
 - 結果的に医薬品開発サイクルの飛躍的な迅速化が進む
- 生成AIの利活用にはデータの標準化・構造化・集約化が必要
 - AIによるアウトプット向上には、インプット側（プロトコール等）の標準化が必要
 - 臨床試験の情報に加えて、作業プロセスも生成AI利活用を前提とした見直しが迫られる
 - 生成AIの基礎が学習データ量であることを踏まえると、日本の課題である施設間データ連携の遅れを解決する試みも重要