

ゲノム情報等を用いたAI創薬ターゲット探索プラットフォームについて

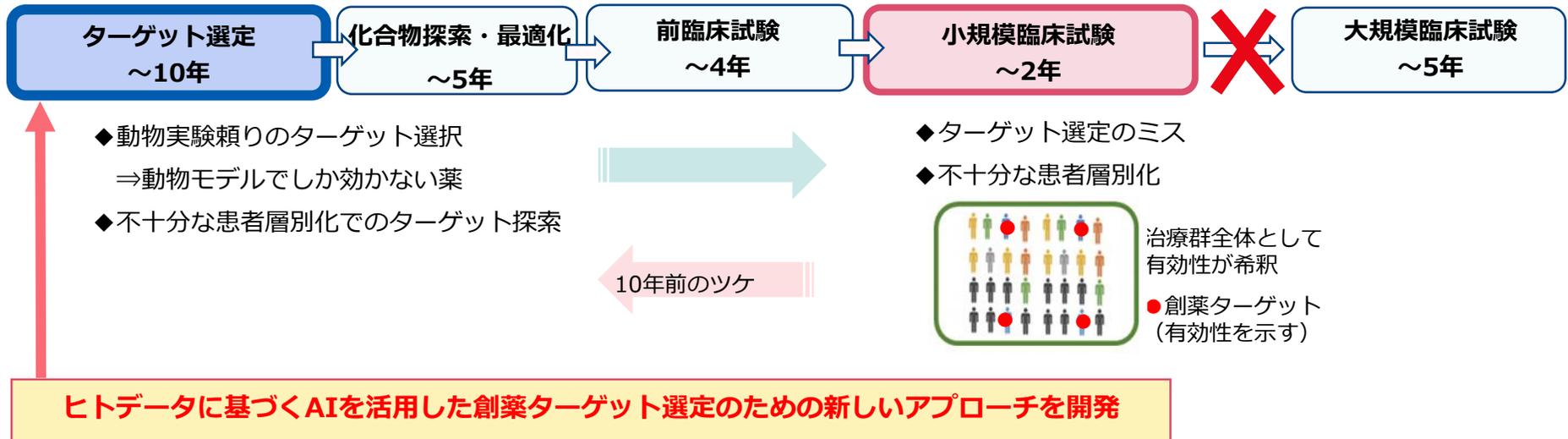
令和4年3月2日

厚生労働省

Ministry of Health, Labour and Welfare of Japan

背景

- ◆ 医薬品開発において、小規模臨床試験（Phase-2）でPOC取得（Proof of Concept）に成功するプロジェクトは30%以下であり、それまでに必要とする期間は平均9年。
- ◆ 開発の失敗には、動物実験頼りの創薬ターゲット選定や不十分な患者層別化によるものが一因と考えられ、**ヒトデータに基づくAIを活用した新たなアプローチ**が有用と考えられる。



◆ 医薬品開発フェーズにおけるAIの活用状況

AI活用が進んでいない領域



- Virtual Screening
- computer-aided drug design
- 構造活性相関 (構造⇔生物活性・毒性・ADME他)



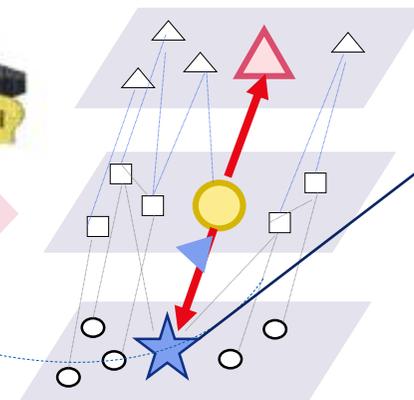
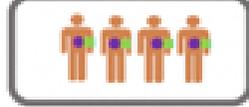
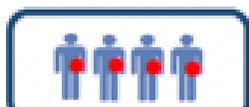
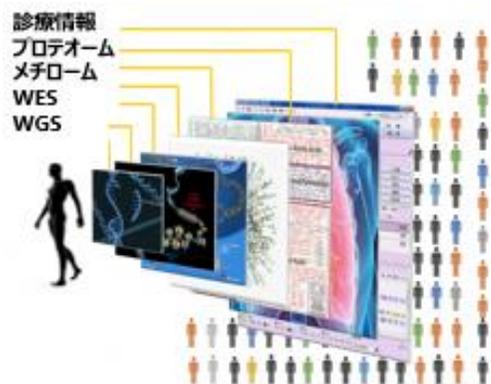
AI活用が進んでいる領域

- AI支援治験患者登録
- バーチャル治験(遠隔モニタリング)
- 自動ケースレポート作成



ゲノム×AIによる創薬ターゲット探索プラットフォームの構築

官民研究開発投資拡大プログラム（PRISM）の枠組みに基づき、平成30年度～令和4年度にかけて、オミックスデータや臨床情報等のヒトデータに基づくデータ駆動的な創薬ターゲット探索手法等を構築。



取組① データ収集

特発性肺線維症及び肺癌患者1500例の臨床情報及びマルチオミックス解析データを収集し、疾患統合データベースを構築

読影所見等の臨床テキスト情報を②の解析に用いるために自然言語処理システムを構築

取組② データ解析

臨床情報及びマルチオミックス解析データを用いて創薬ターゲット／患者層別化バイオマーカー等を探索できるAIを開発し、その性能を実証

取組③ 結果解釈

②で得られた分子に関する文献調査の効率化を図るため、特定した創薬ターゲット候補の分子機能や分子間作用を科学論文等から構築した知識ベースを基に推定する技術を開発

取組④ オープンプラットフォーム

事業成果を広く健康・医療領域の研究開発に展開するために、成果物をオープンプラットフォーム化

今後の方向性（案）

- 「全ゲノム解析等実行計画*」においては全ゲノム情報に加え、オミックスデータや診療情報等を収集することとしており、これらの集積されたデータを用いて、AI活用を含む高度な横断的解析の推進が厚生労働科学研究班**より提言されている。
- 「全ゲノム解析等実行計画」とは別に、官民研究開発投資拡大プログラム（PRISM）の枠組みに基づき、平成30年度～令和4年度にかけて、全ゲノム情報などのオミックスデータや診療情報等を収集し、それらのデータをもとに患者層別マーカーと創薬ターゲットを選定するAIプラットフォームを開発中。
- これら二つのプロジェクトが連携することで、より効率的に、新たな創薬ターゲットを創出することが期待できる。
- 今後、「全ゲノム解析等実行計画」における全ゲノムデータ等の利活用体系の一つの選択肢として、PRISMにおいて構築しつつある創薬ターゲット探索プラットフォームの利活用について検討してはどうか。

* 全ゲノム解析等実行計画（第1版）令和元年12月20日策定（厚生労働省）

** 「全ゲノム解析を基盤としたがんゲノム医療の実装に向けた患者還元、解析・データセンター、ELSI 体制構築についての研究」班（厚労科研中釜班）

参考資料

ひと、暮らし、みらいのために



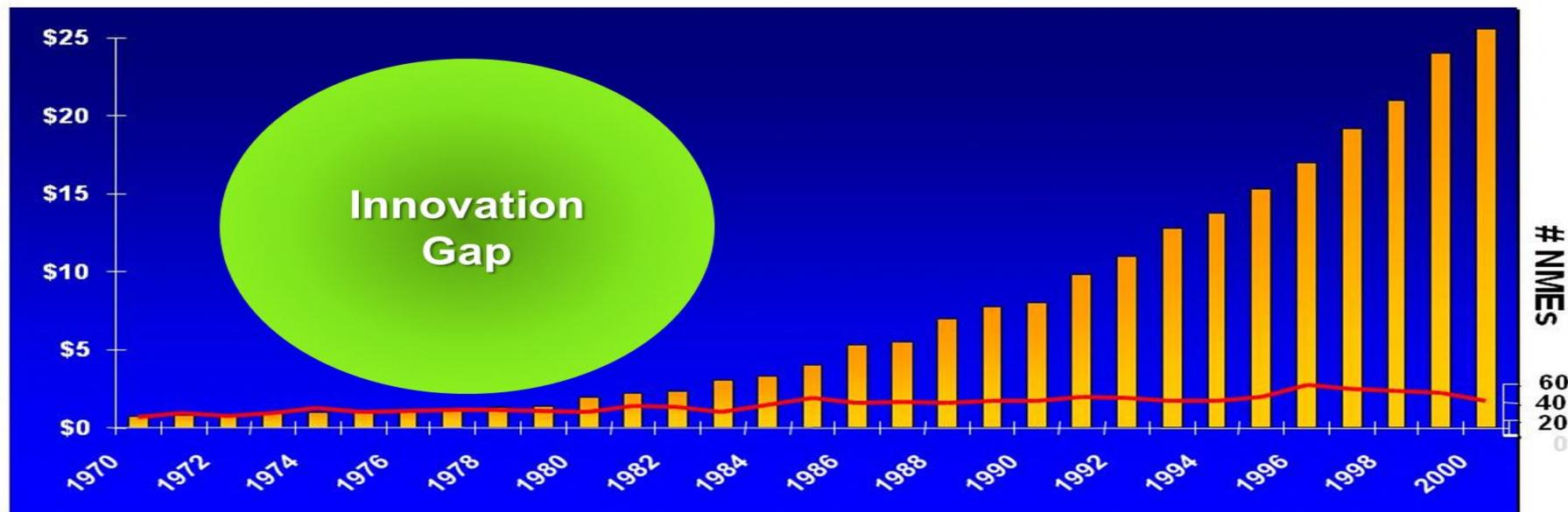
厚生労働省
Ministry of Health, Labour and Welfare

本事業の背景 —新薬メーカーが直面する世界的課題—

- ◆ **R & D生産性の低下**
 - 新薬開発の対象が、原因が未解明の複雑な疾患にシフト
 - 成功確率の低下
 - 研究開発コストの高騰
- ◆ 薬事規制の強化
- ◆ 薬価引き下げ圧力
- ◆ ジェネリック医薬品との競合

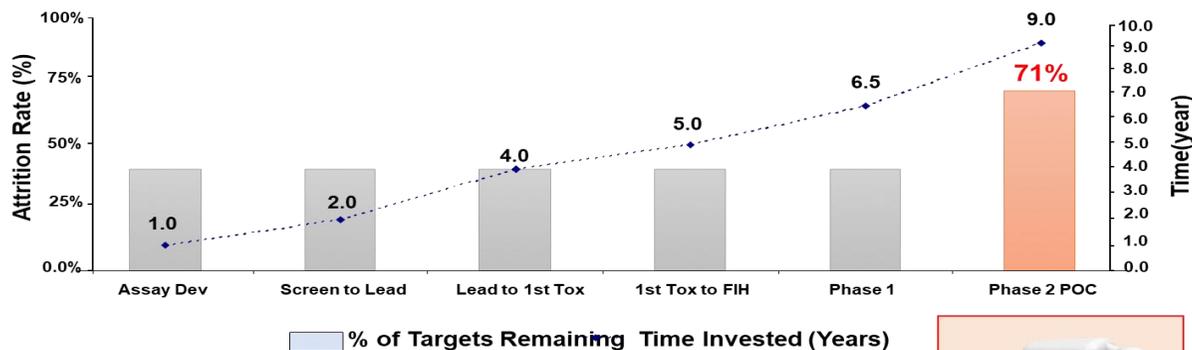
製薬産業界のイノベーションギャップ^o

Total R&D Investment (\$ Billions)



本事業の背景 — 医薬品開発の最大の隘路 —

POC取得（Phase2）に成功するプロジェクトは30%程度
POC取得に要する期間は平均9年

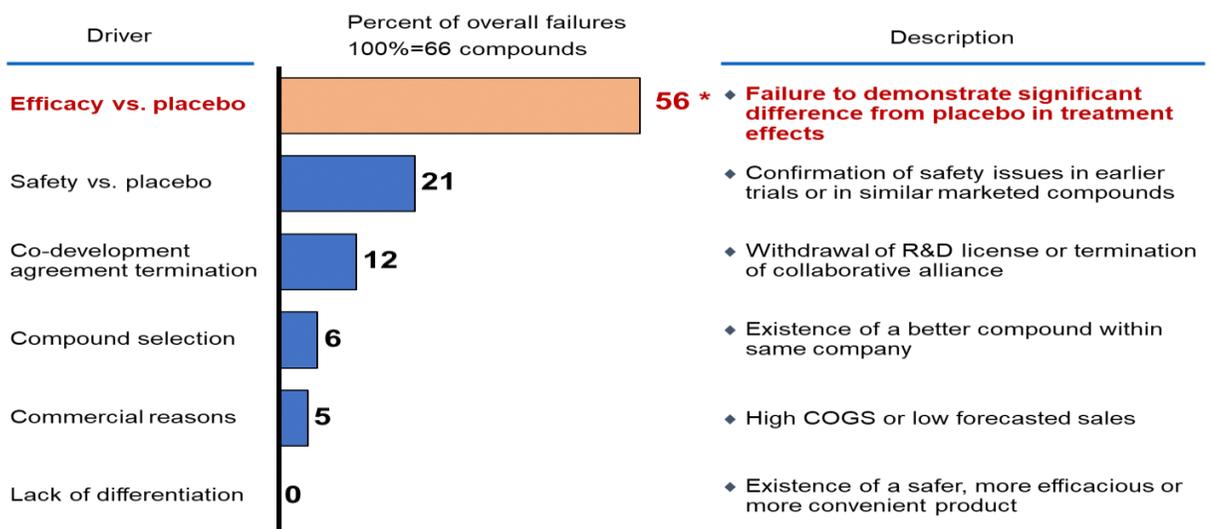


Phase-2でのPOC取得 (Proof of Concept)の失敗率が最も高い

Assumptions: 1st 5 stages have ~ 60% success rate each, and Phase 2 has ~ 30% success rate. Times adapted from various sources.



POC = 創薬開始時に想定した有用性の証明



POC取得失敗の原因で最も多いのが「**有意な薬効が得られなかった**」

* Only 1 indication RIP out of 37 indication RIPs by efficacy vs. placebo had an established MOA
Source: EvaluatePharma, Pharmaprojects; Factiva, Literature Search; team analysis

官民研究開発投資拡大プログラム（PRISM）概要

- 官民研究開発投資拡大プログラム（PRISM）は、総合科学技術・イノベーション会議（CSTI）の司令塔機能を強化するために、平成30年度に創設（100億円）。

【目的】

民間研究開発投資誘発効果の高い領域又は財政支出の効率化に資する領域への各府省庁施策の誘導を図ることを目的とする。

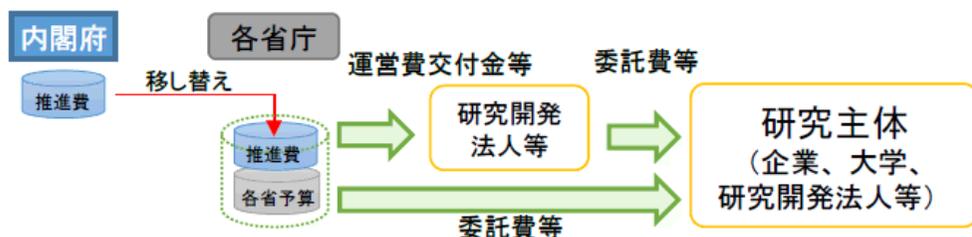
【事業概要】

- ガバニングボードは、領域を設定し、予算の配分等に強い権限をもった領域統括の下、統合イノベーション戦略に基づく各種戦略の実現に必要な施策を、トップダウンで決定。対象施策に係る研究開発の加速、新規研究開発課題の前倒し等に必要となる経費を、内閣府より追加配分。＜研究開発型：平成30年度より実施＞

【令和3年度領域：AI技術、建設・インフラ維持管理／防災・減災技術、バイオ技術、量子技術】

- 中長期的に官民研究開発投資の拡大を図るため、令和元年度から、国立大学における民間資金獲得を推進する事業、令和2年度から、スタートアップ・エコシステム拠点形成による創業環境整備を推進してスタートアップを支援する事業を実施し、令和3年度からは、科学技術・イノベーション創出の活性化に関する法律に基づく新SBIR制度における省庁連携を加速させる事業、社会課題解決や国際市場獲得等を促進する標準活用施策の加速化支援をする事業を実施予定。＜システム改革型＞

資金の流れ



期待される効果

- イノベーション創出につながる官民研究開発投資の拡大。
- 民間資金・寄付金など外部資金を拡大できる経営基盤の形成による、大学等に対する企業の投資額の増大。新SBIR制度の加速と、スタートアップ・エコシステム拠点の形成による、スタートアップへの投資額の増大。標準活用の加速化による社会課題解決や国際市場獲得。

PRISM事業の現状

取組① データ収集



全ゲノムを含む多層化オミックスデータ及び臨床情報を収集

特発性肺線維症 (IPF)

世界初の臨床情報が紐づいたオミックスデータのデータベースを構築 (R3に1,308症例見込み)。

肺がん患者

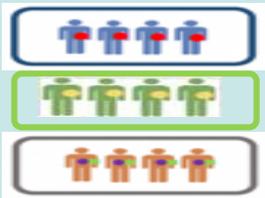
米国のTCGA (1,089症例) を上回る1,624症例のデータベースを構築。

診療テキスト構造化システム

IPF、肺がんについて、オントロジー、アノテーション・コーパス、知識ベースの整備を行い、医療分野テキストを高度に解析・構造化する言語・知識処理基盤を構築。

取組② データ解析

対象疾患患者データから創薬ターゲットを探索する技術を複数開発



患者層別化 診療情報及びオミックスデータという異種のデータを入力し、症状等に密接に関連した生体分子を特定。

バーチャル創薬・臨床試験 新薬に対して、マウス遺伝子発現からヒトでのアウトカムを予測。

疾患特異的サブネットワーク 公共DBから構築した遺伝子ネットワークに患者データをマッピングして層別化。

創薬ターゲットを共有する疾患予測 遺伝子発現データから算出した疾患ごとの類似度を用いた創薬標的探索。

深層学習を使った解析システム 深層学習技術を活用してオミックスデータをマルチモーダルに解析。

取組③ 結果解釈

文献等の既存知識の調査を効率化するためのツールを構築又は開発中

TargetMine

IPFと肺がんのデータを追加し、結果解釈に利用。

産総研のシステム

文献から知識ベースを構築し、そこから分子機能や分子間作用を推定するシステムを開発中。

取組①～③による成果

創薬ターゲット候補

IPFと肺がんでこれまで報告されていない分子/パスウェイを見出した！

特発性肺線維症 (IPF)

IPFでの関与が報告されていない分子/パスウェイを見出した (用途特許出願済み1件、出願予定1件)。ヒト肺疾患部での発現を確認 (免疫染色) したほか、家族性IPFの遺伝子変異導入マウスでの実験や患者でのメタボローム解析などを実施し、検証中。

肺がん患者

Pan-negative肺がんの創薬ターゲットを探索し、関与が報告されていない重要分子1つを見出した。今回同定された分子は、全ゲノムデータ・ChIP-seqデータ・RNA-seqデータを統合的に解析しない限り同定は困難であった (全ゲノム解析を中核とした多層化オミックスデータをマルチモーダルに解析する重要性を立証)。現在PDXマウスで検証中。

PRISM事業のガバナンス体制

- : 厚労省
- : 文科省
- : 経産省

◆ マネジメント体制

(PD・全体とりまとめ)
榎林 陽一
 神戸大特命教授

連携

プロジェクトマネジメント
 オフィス
 ・ プロジェクトの推進、
 管理、調整等

(理研-PD)
上田 修功
 理研AIP副センター長

【省庁担当】
 厚生労働省
 大臣官房厚生科学課長
 文科科学省
 研究振興局
 参事官(情報担当)

◆ 対象施策実施体制

医薬基盤・健康・栄養研究所 夏目やよい
 ・ 肺がん及びIPF患者情報収集
 ・ IPF統合DB構築
 ・ 創薬ターゲット推定アルゴリズムの開発
 ・ 統合データウェアハウス拡充
 ・ 創薬ターゲット候補の検証

理化学研究所 上田 修功
 ・ 高度AI構築のための指導・支援

- 産業技術総合研究所 高村 大也
 ・ 情報自動抽出システム ・ 生体分子相互作用推論システム
- 神奈川県立循環器・呼吸器病センター 小倉 高志
 ・ IPFを含む間質性肺炎患者情報の収集
- 大阪大学 熊ノ郷 淳
 ・ 層別化AIの精度評価に向けたIPF患者の解析
- 京都大学 奥野 恭史
 ・ 創薬ターゲット推定アルゴリズムの開発
- Karydo TherapeutiX 佐藤 匠徳
 ・ 創薬ターゲット推定アルゴリズムの開発
- 北海道大学 西村 紳一郎
 ・ オミックス解析結果の解釈精度向上のための糖鎖解析
- 国立がん研究センター 浜本 隆二
 ・ 肺がん患者情報収集 ・ 肺がん統合DB構築
 ・ 創薬ターゲット推定アルゴリズムの開発
- Preferred Networks・岡野原 大輔
 ・ 肺がん統合DB構築 ・ 創薬ターゲット推定アルゴリズムの開発
- 産業技術総合研究所 瀬々 潤
 ・ 肺がん統合DB構築 ・ 創薬ターゲット推定アルゴリズムの開発
- 京都大学 黒橋 禎夫
 ・ 医療テキスト構造化のための言語・知識処理基盤の構築
- 奈良先端科学技術大学院大学・荒牧 英治
 ・ オントロジーとアノテーション・コーパスの構築
- 大阪大学・荒瀬 由紀
 ・ 医療テキストからの情報抽出
- お茶の水女子大学・戸次 大介
 ・ 表現計算モデルの構築
- 九州工業大学 山西 芳裕
 ・ 創薬ターゲット候補と疾患の関連を予測するシステム開発
- 理化学研究所 田部井 靖生
 ・ アルゴリズムの開発
- 大分大学 加藤 明良
 ・ 層別化AIの精度評価に向けた疾患モデルの解析等
- 理化学研究所 瀧川 一学、田部井 靖生
 ・ アルゴリズムの開発