様式３

創薬ターゲット探索プラットフォームの構築

新薬創出を加速する症例データベースの構築・拡充/創薬ターゲット推定アルゴリズムの開発

官民研究開発投資拡大プログラム（PRISM）

「ＡＩ技術領域」

令和４年度成果

令和５年３月

厚生労働省・文部科学省

資料1 「創薬ターゲット探索プラットフォームの構築」の概要

アドオン：581百万円（厚生労働省・文部科学省）/元施策有/PRISM事業・継続予定

課題と目標

■課題：動物実験頼りで創薬ターゲットを選定する従来の研究スキームが医薬品開発の最大のボトルネックとなっている第2相試験（Phase 2）での成功率低迷の要因となっているため、製薬産業界では、患者の診療データ等とAIを活用して患者を層別化し、データ駆動的に適切な創薬ターゲットを探索する新たなプラットフォームの開発が新薬創出のゲームチェンジャーとして切望されている。

■目標：特発性肺線維症 (IPF) 及び肺がんを対象に、診療情報と網羅的生体分子情報から患者を層別化し、創薬ターゲットをデータ駆動的に探索するプラットフォームの構築に必要となる多岐にわたるAI群を開発する。さらに、 IPF及び肺がんの新たな創薬ターゲットをそれぞれ1個ずつ特定し、企業導出等により産業界での実創薬に展開する。

「創薬ターゲット探索プラットフォームの構築」の概要

■元施策：

厚生労働省（医薬基盤・健康・栄養研究所）では、平成29年度より、「新薬創出を加速する人工知能の開発」事業において、 IPFを対象に、診療情報と網羅的生体分子情報を収納した疾患統合データベースの構築と新たな創薬ターゲット探索手法の開発を開始。令和4年度は「健康・医療ビッグデータ活用システムの構築」を元施策としている。（令和4年度199百万円）

■PRISMで実施する理由：

- 日本が強みを持つ創薬領域での研究開発を強化するプロジェクトであり、創薬ターゲット探索に焦点をあてたAI開発は世界に先駆けた取り組みである。
- 本施策により、医薬品開発コストの大幅な削減と個別化医療の進展がもたらされ、薬剤費及び医療費の低減が期待される。
- 本施策により、有効な治療薬がない肺がん、IPFについて、アンメットメディカルニーズを満たす新たな医薬品の創出に繋がることが期待される。
- 事業成果のオープンプラットフォーム化や導出等により企業と公的機関（公立病院、大学、国研等）とのAI開発等での連携強化を目標としている。

■テーマの全体像 ：

IPFと肺がんを対象に、日本トップクラスの医療機関との連携のもと、質の高い診療情報と網羅的生体分子情報をそれぞれ1500症例以上収集し、世界最大規模の疾患統合データベースを構築する。さらに、これらのデータや科学論文、既存データベース等を活用して、データ駆動的に患者を層別化し、創薬ターゲットを探索する世界最先端のAI群とデータウェアハウスを開発する。医薬健栄研、理研、大学・国研など15機関が参画。

出口戦略

- 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、疾患データベース、アルゴリズム、研究スキーム等をオープンプラットフォーム化する。オープンプラットフォーム化にあたっては、関係各所との調整のうえ、持続的かつ自立化が可能な事業運営を目指す。
- オープンプラットフォーム化になじまない創薬ターゲット及び本事業の実施過程で派生的に創出された技術・成果物については民間企業への導出を図る。

民間研究開発投資誘発効果等

- 民間からの貢献額：令和4年度実績として民間企業52社（製薬企業：14社、IT企業：16社、化学・食品企業：12社、その他：10社）と総額4.9億円相当の共同研究を行った。（内訳）：共同研究費（1.1億円）、人件費（3.8億円）、資機材（0.04億円相当）
- 民間企業への導出（令和元年度）：AI搭載ITプラットフォーム（経済効果50億円：導入企業による試算）
  （令和３年度）：医師がAI技術を開発できる研究基盤システム「AI開発支援プラットフォーム」の共同開発

資料2　「創薬ターゲット探索プラットフォームの構築」の概要

テーマの全体像・説明

本テーマでは、新薬開発の最大のボトルネックであるPhase 2での成功確率を上げるため、AIや大規模知識ベース等の最先端情報科学を実装した革新的な“データ駆動型の創薬標的探索プラットフォーム”を構築する。

PRISMによる加速・拡充

厚労省、文科省が密接な連携体制を構築し、

- 製薬業界が抱える課題について、官民共同でAIを活用した創薬ターゲット探索の仕組みを構築し、効率的な創薬システムの実現を加速する。
- 対象疾患として難病であるIPFに加えて、がんの死因第1位の肺がんも対象として拡充する。

「創薬ターゲット探索プラットフォームの構築」の概要

本施策では、① 診療情報及びマルチオミックス情報を収集し、それらを構造化した疾患統合データベースを構築する。② 疾患統合データベース及びAIを活用して患者を層別化し、診療情報と密接に関連した生体分子を探索する。③ AIにより探索された生体分子の分子機能や分子間相互作用を、既存知識を利用して推定することで、創薬ターゲット候補を決める。この技術を民間企業との連携を通じて社会実装させることで民間研究開発投資を促進する。

厚労省 元施策

「健康・医療ビッグデータ活用システムの構築」
《令和4年度予算》
199百万円

PRISMによる加速・拡充

「創薬ターゲット探索プラットフォームの構築」
《令和4年度予算》
厚労省：574百万円
文科省：　7百万円

① データ収集

- 診療情報と患者検体のマルチオミックス情報の収集

(1) IPF DBの構築
IPFマルチオミックス解析
医薬健栄研

② データ解析

- AIを活用し、診療情報と紐づいた分子を探索

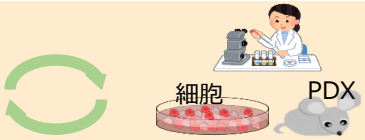
(2) 創薬ターゲット推定
アルゴリズム開発
医薬健栄研

③ 結果解釈

- 生体での分子機能や分子間相互作用を推定

(3) データウェアハウスを利用した分子機能等の推定
医薬健栄研

《 創薬ターゲット候補の検証 》



検証結果
フィードバック

AIによって示された
創薬ターゲット候補の検証

《 DB構築の加速・拡充 》

(4) 肺がんマルチオミックス解析
医薬健栄研

(5) 肺がんDBの構築
国がん

(6) カルテ・患者日誌等の自然言語処理技術開発
京大

《 解析の高度化 》

(7) AI 要素技術の研究開発
医薬健栄研, Karydo,
理研, 京大, 九工大

(8) 創薬ターゲット推定
アルゴリズム開発の指導
理研

(9) 創薬ターゲット推定
アルゴリズム開発
国がん

《 検証能力の効率化 》

(10) 高度な統合解析プラットフォームの構築
医薬健栄研, 北大, 阪大

(11) 情報自動抽出及び推論システムの構築
産総研

(12) 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
九工大

民間企業との連携

（民間企業）

- 製薬企業：14社
- IT企業　：16社
- 化学・食品企業：12社
- その他：10社

④ オープンプラットフォーム化の加速

(1～12) で作られたDB及びAIを含むプログラム等をアカデミア・企業等に公開する仕組み作り
医薬健栄研

資料3　「創薬ターゲット探索プラットフォームの構築」の目標達成状況

〇施策全体の目標
　データ駆動的に患者を層別化するなど、創薬ターゲットを探索するAI群の開発及びIPFと肺がんの新たな創薬ターゲットをそれぞれ1個ずつ特定し、共同研究や企業導出等により産業界での実創薬に展開する。

事業全体目標	当年度目標	目標の達成状況
①（データ収集） • IPF及び肺がんで**1,500症例**の診療情報が紐づいたマルチオミックスデータの収集。	①（データ収集） • IPFデータベース構築：累計1,500症例の達成 • 肺がんデータベース構築：オミックスデータ欠損値補完	• IPF：今年度に**目標達成済み**（累計1,503症例）。肺がん：令和元年度に**目標達成済み**（累計1,714症例）。令和2年度よりpan-negative肺がんを中心に欠損値を補填した。
IPF及び肺がんから **各1個の創薬ターゲットの特定** ②（データ解析） • 患者の層別化、疾患や症状に特徴的な分子の抽出技術の完成。 ③（結果解釈） • 文献からの知識の抽出/知識表現及び推論技術の完成。	②（データ解析） • 複数の異なるAIアルゴリズムを試行し、創薬ターゲット候補を追加提案 ③（結果解釈） • 創薬ターゲット候補の確からしさを検証する • 文献からの情報自動抽出及び推論技術の完成	• 創薬ターゲットとして**IPFでタンパク質A及び複数のチロシンキナーゼ**を、pan-negative**肺がんでタンパク質B**を見出した。いずれも動物実験（IPF：プレオマイシンモデル及び家族性IPFモデルマウス、肺がん：PDXマウス）にて効果を確認し、**目標達成済み**。 • IPFを含む進行性線維化を伴う間質性肺疾患（PF-ILD）の**バイオマーカー**を見出した。従来IPFのバイオマーカーとされてきたKL-6より優れていた。 • 創薬ターゲットについての特許2件を出願済み。追加で特許1件を出願予定。 • これまでに開発した手法を用いて、患者データの解析を続行するとともに、アルゴリズムの改良を実施した。 • IPFでの関与が報告されていない分子/パスウェイについて引き続き検証を行っている。 • 英語文献から取り出した情報の分類（疾患名なのか、薬剤名なのか、生体物質名なのか等）及び情報間の関係性（原因なのか、結果なのか、促進なのか、抑制なのか等）の抽出を行い、疾患に関連する事象や物資の関係性を示すネットワーク（知識ベース）を作るシステムの構築を**完了**した。
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、疾患データベース、アルゴリズム、研究スキーム等をオープンプラットフォーム化する。	④（オープンプラットフォーム） • 海外運用の開始 • 国内運用の継続 • 国内ユーザーの声を反映したシステムの改良	• 個人情報保護法の規定上、国立研究開発法人が匿名加工情報を作成/提供できるかについて、法律専門家の間で意見が割れているために**オープンプラットフォームでの匿名加工情報の提供は当面見送らざるを得ないと判断した**。 • **匿名加工情報を除き、**事業で開発したAIのみ**オープンプラットフォームの運用を開始**。 • **オープンプラットフォームの共同研究領域は順調**で、既に昨年度報告したGAPFREE4事業及びLINCに加え、①A社（他事業での収集データ+AIの活用）と②B社（新規収集データ+AIの活用）、③C社（本事業で収集したデータを用いた解析）は共同研究契約を締結し、④D社（本事業で収集したデータを用いた解析）は、秘密保持契約を締結し、共同研究について協議を始めた。

資料4　「創薬ターゲット探索プラットフォームの構築」の成果

○ データ収集

- IPFは累計1,503症例の診療情報が紐づいたオミックスデータを収集し、目標を達成した。
- 肺がん症例数は世界最大規模（1,714症例）を達成済み。令和2年度よりpan-negative肺がんを中心に欠損値を補填した。

○ データ解析

- 診療データと紐づくDNA、RNA、たんぱく質等のデータからAIを使ってデータ駆動的に創薬ターゲットを効率的に探索する新しい研究スキームを確立した。
- 創薬ターゲットとして、IPFでタンパク質A及び複数のチロシンキナーゼを、pan-negative肺がんでタンパク質Bを見出した。いずれも動物実験にて効果を確認した（結果解釈の項参照）。また、IPFを含む進行性線維化を伴う間質性肺疾患（PF-ILD）のバイオマーカーを見出し、従来IPFのバイオマーカーとされてきたKL-6より優れていることを確認した。
- IPFについて、新たに開発した手法により本事業で収集した患者情報を解析中。また、上記のタンパク質以外に医薬健栄研が本事業で開発した層別化AI（SubsetBinding）を用いた解析により複数の新規及び既知の創薬ターゲット候補分子が見出された。これらについては、ヒト細胞や動物モデルにより検証中。
- 読影所見等の臨床/患者テキストからの医療情報抽出システムについて、症例テキストを追加し、大幅に精度を改善した。
- 42の解析プログラム（自然言語処理プログラムを含む）を作成し、内18のプログラムをオープンプラットフォームで公開した。

○ 結果解釈

- IPFで見出された創薬ターゲット候補分子について、患者の血清/尿中代謝物の網羅的測定、in vitroの上皮間葉転換試験系及びin vivoのブレオマイシンによるIPFモデルマウスで検証を行い、候補分子が関与している可能性を支持する結果が得られた。
- タンパク質Aの機能を阻害する薬剤をブレオマイシン肺線維症マウス及びSP-Cノックイン肺線維症モデルに投与したところ、肺組織のコラーゲン沈着が有意に抑制された。タンパク質Bに関しては、過剰発現している肺がん症例のPDXモデルにタンパク質Bの機能を阻害する薬剤を投与し、用量依存的に有意な腫瘍縮小効果が示された。タンパク質Bの機能を阻害する薬剤の人への投与を目指して研究を発展させていく予定である。
  タンパク質A及びBのいずれも、データ駆動的に創薬ターゲットを見出したことを投稿予定。
- 英語文献から取り出した情報の分類（疾患名なのか、薬剤名なのか、生体物質名なのか等）及び情報間の関係性（原因なのか、結果なのか、促進なのか、抑制なのか等）の抽出を行い、疾患に関連する事象や物資の関係性を示すネットワーク（知識ベース）を作るシステムの構築が完了した。
- オミックス解析結果の解釈精緻化を目的に、各種臓器/組織及び血中エクソソームの糖鎖解析を行い、世界初の「マウス組織グライコームアトラス」に関する論文を投稿した。これは臓器/組織がどのような糖鎖で修飾されているかをカタログ化したものであり、血中エクソソームの糖鎖と照らし合わせてどこの組織に由来するかを推測するための一歩となった。さらに、「ヒト組織グライコームアトラス」の作成に向け徳島大学との共同研究に着手した。

○ オープンプラットフォーム

- オープンプラットフォームのレベル2で匿名加工情報を搭載することについて、個人情報保護法上の懸念が生じているため、オープンプラットフォームでの匿名加工情報の提供は当面見送らざるを得ないと判断した。そのため、オープンプラットフォームは、匿名加工情報を除き、本事業で開発した汎用的で有用性が高い18のAIを搭載して運用を開始した。
- オープンプラットフォームのレベル3（共同研究領域）での利活用は順調に進捗しており、既に昨年度報告したGAPFREE4事業（製薬企業7社）及びLINC（製薬及びIT企業7社）に加え、①A社（他事業での収集データ+AIの活用）、②B社（新規収集データ+AIの活用）、③C社（本事業で収集したデータを用いた解析）は共同研究契約を締結し、④D社（本事業で収集したデータを用いた解析）は秘密保持契約を締結し、共同研究についての協議を開始した。

資料5　「創薬ターゲット探索プラットフォームの構築」の民間からの貢献及び出口の実績

〇民間からの貢献額：平成30年度から令和4年度の5年間で22億円相当

① 平成30年度　5.2億円（内訳）共同研究等29社：共同研究費（3億円）、人件費（2.0億円）、資機材（約0.2億円相当）
② 令和元年度　4.3億円（内訳）共同研究等32社（製薬企業20社を含む）　：共同研究費（1.7億円）、人件費（2.4億円）、資機材（約0.2億円相当）
③ 令和2年度　4.5億円（内訳）共同研究等44社（製薬企業24社を含む）：共同研究費（1.0億円）、人件費（3.4億円）、資機材（約0.1億円相当）
④ 令和3年度　4.2億円（内訳）共同研究等43社（製薬企業17社を含む）：共同研究費（0.7億円）、人件費（3.5億円）
⑤ 令和4年度　4.9億円（内訳）共同研究等52社（製薬企業14社を含む）：共同研究費（1.1億円）、人件費（3.8億円）、資機材（0.04億円相当）

当年度当初見込み	当年度実績
新薬創出を加速する症例データベースの構築・拡充（厚労省）	新薬創出を加速する症例データベースの構築・拡充（厚労省）
① 創薬ターゲット探索に関する共同研究　人件費（53百万円、28社） ② in silico創薬に関する共同研究　4件　人件費（65百万円） ③ バイオマーカー関連の共同研究　2件　研究費（1.5百万円）、人件費（6百万円） ④ 創薬関連の共同研究　3件　研究費（30百万円）、人件費（70百万円） ⑤ 創薬関連の共同研究　2件　研究費（8.5百万円）、人件費（5百万円） ⑥ 自然言語処理に関する共同研究　3件　研究費（6.7百万円）、人件費（12百万円） ⑦ インフォマティクスに関する共同研究　5件　研究費（21百万円）、人件費（40百万円） ⑧ がん関係の共同研究費　4件　研究費（15百万円）、人件費（130百万円） ---------- 計：共同研究等47社（製薬企業15社を含む）：共同研究費（0.8億円）、人件費（3.8億円）	① 創薬ターゲット探索に関する共同研究　人件費（53百万円、28社） ② in silico創薬に関する共同研究　4件　研究費（7.8百万円）、人件費（48百万円） ③ バイオマーカー関連の共同研究　1件　研究費（1百万円）、人件費（2百万円） ④ 創薬関連の共同研究　8件　研究費（45.3百万円）、人件費（53.5百万円） ⑤ 健康関連の共同研究　8件　研究費（25.6百万円）、人件費（72百万円） ⑥ 自然言語処理に関する共同研究　4件　研究費（13百万円）、人件費（14百万円）、資機材（4百万円） ⑦ インフォマティクスに関する共同研究　3件　研究費（11百万円）、人件費（26百万円） ⑧ がん関係の共同研究費　3件　研究費（5百万円）、人件費（110百万円） ---------- 計：共同研究等52社（製薬企業14社を含む）：共同研究費（1.1億円）、人件費（3.8億円）、資機材（0.04億円）

〇出口戦略

- 優れた創薬ターゲットを提供することで、民間製薬企業の医薬品開発を喚起し、研究開発投資誘発効果を期待。（承認薬１剤当たりの研究開発費約3千億円）
- 医療・ヘルスケア領域全般に展開できる成果を提供することで、民間企業の研究開発投資誘発効果が期待できる。
  例）医療情報システム市場：電子カルテや医用画像関連システム等の医療情報システム全体の国内市場規模は約6,000億円、電子カルテだけで約4,000億円と言われており（JAHISの調査研究）、本事業の成果導出により研究開発投資効果が期待できる。

当年度当初見込み	当年度実績
• 民間企業からの共同研究等に関する相談・調整等：10社 （創薬ターゲット探索、薬効予測、読影所見の自動生成、AI技術を用いたがん医療プラットフォームの製品化及び副作用情報の自動抽出等の自然言語処理技術など）	• 民間企業からの共同研究等に関する相談・調整等：10社 （創薬ターゲット探索、薬効予測、読影所見の自動生成、AI技術を用いたがん医療プラットフォームの製品化、副作用情報の自動抽出等の自然言語処理技術及びヘルスデータの解析など）