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」

【継続事業】

施策説明資料

令和4年3月 厚生労働省·文部科学省

「創薬ターゲット探索プラットフォームの構築」の全体像

背景

Phase-2でPOC取得(Proof of Concept)に成功するプロジェクトは30%程度、それまでに必要とする期間は平均9年。

➡不十分な患者層別化と創薬ターゲット選定のミスが原因 ➡AIを活用した新たなアプローチを開発する

Game Changer (AI) の開発に大きな期待

AI活用が進んでいる領域



- 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



ターゲット選定(仮説創出) ~10年

探索 · 最適化 ~5年

前臨床試験(安全性等) ~4年

仮説検証(小規模治験) ~2年

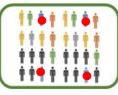
有用性確認(大規模治験) ~5年

- ◆動物実験頼りのターゲット選択➡動物モデルでしか効かない薬
- ◆不十分な患者層別化でのターゲット探索

10年前のツケ

◆不十分な患者層別化

◆ターゲット選定のミス



治療群全体として 有効性が希釈

●創薬ターゲット (有効性を示す)

PRISMでAI技術を活用し、創薬ターゲット選定のための新しいアプローチを開発

PRISMでの取組の全体像

取組① データ収集 (目標)

IPF及び肺がん患者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(疾患

DB) を構築

(創出される成果物)

新たな構造化電子カルテモデル

日本初の症例報告自動作成 システム実装モデル

成果初のIPF・肺がん患者層別化 医療情報統合データベース

医療テキスト解析システム (診療録からのサマリ自動生成、病理 所見と診断の整合性チェックなど)

画像自動アノテーションシステム

取組② データ解析 (目標)

創薬ターゲット/患者層別化バイオ マーカーの探索に必要なAIを開発し、 創薬ターゲット候補を探索

(創出される成果物)

創薬に向けた新たなAI技術 (新規解析手法·深層学習)

新たな患者層別化バイオマーカー (IPF・肺がん)

取組③ 結果解釈 (目標)

候補を文献等から機能等を推定

創薬ターゲット/患者層別化バイオ マーカーの探索に必要な知識ベースを 開発し、AIが推測した創薬ターゲット

(創出される成果物)

科学論文から知識を自動検索・ 抽出する新たな技術

科学論文から自動抽出した知識 から生体分子間相互作用を自動 推論する新たな技術

> 世界初のIPF・肺がんの 大規模知識ベース

(目標)

IPFと肺がんの創薬ターゲットを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

創薬ターゲット

(創出される成果物)

創薬ターゲット

取組4 オープンプラット フォーム

(目標)

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

(創出される成果物)

オープンプラットフォーム

IPF・肺がんの創薬に展開できる成果

他疾患の新薬創出及びヘルスケア 領域全般に展開活用できる成果

資料2 元施策「新薬創出を加速する人工知能の開発」の概要

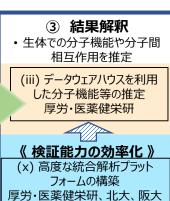
テーマの全体像・説明

- 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。
- 本事業では、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結集し 革新的なデータ駆動型の創薬ターゲット探索プラットフォームを構築する。

「新薬創出を加速する人工知能の開発」の概要







(xi) 情報自動抽出及び

推論システムの構築

経産・産総研

(xii) 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

文科·JST/九工大

民間企業との連携 (民間企業) • 製薬企業:17社

検証結果

フィードバック

• IT企業 : 13社 · 化学·食品企業: 9社

《 創薬ターゲット候補の検証 》

0.01 000

(xiii) AIによって示された

創薬ターゲット候補の検証

• その他: 4社

(vi) カルテ・患者日誌等の自 然言語処理技術開発 文科·JST/京大

文科·理研

(ix) 創薬ターゲット推定 アルゴリズム開発 文科・JST/国がん

4 オープンプラットフォーム化の加速

(i~xii)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 厚労・医薬健栄研

ロードマップ

文科省:

データベースの拡充/データ収集(対象疾患: IPF、肺がん) アドオン施策 解析技術の高度化・加速化/オープンプラットフォームの構築 R2年度 R3年度 R1年度 H30年度

○データ解析 ○結果解釈 ○仮説検証

○データ収集 ○データ収集 ○データ解析

元施策 ○データ収集

7百万円

○データ解析 ○結果解釈 ○結果解釈 ○オープンプラットフォーム

(協定締結、調査)

○データ収集 ○データ解析 ○結果解釈

○オープンプラットフォーム (調査、制度設計)

○データ収集 ○データ解析

○結果解釈 ○オープンプラットフォーム (ルール策定、運用開始)

対象疾患: IPF

R4年度

システム

の完成

○データ収集 ○データ解析

○結果解釈

・疾患統合データベース (IP、肺がんともに 累計約1500例)

AIや大規模知識ベース等の最先端情報

科学を結集し革新的なデータ駆動型の創

薬ターゲット探索プラットフォームを構築する

・情報の収集・構造化のための処理技術

事業終了時の成果イメージ

・新たな仮説の推論システム ・医療情報を医療従事者以外に公開し

利用してもらう新たなモデルケース

○オープンプラットフォーム

<u>資料3-1 「創薬ターゲット探索プラットフォームの構築」の概要①</u>

アドオン 厚生労働省:574百万円、文部科学省:7百万円/元施策 厚生労働省:291百万円

背景·現状

- ◆ 背景
- 現在の新薬開発は、複数の機能分子や環境要因が関与する難病・稀少病等の複雑な疾患へとシフトしている。そのため、従来の遺伝子解析や動物実験 頼みのアプローチでは新たな創薬ターゲットを見出すことが困難であり、「創薬ターゲットの枯渇」が産業界の深刻な問題となっている。
- 最近では、ゲノムから臨床情報までの疾患情報をAIで解析し、創薬ターゲットと患者層別化バイオマーカーを探索する新たなアプローチに「ゲームチェンジャー | として大きな期待が寄せられ、米国ではメガファーマや情報系ベンチャーによる熾烈な開発競争が繰り広げられている。 • 日本でも製薬企業等の意欲は高いが、企業だけで十分な臨床情報収集やデータ解析体制を整備することが困難なことから、国による先駆的取組に大きな
- 期待が寄せられている。 厚生労働省では、平成29年4月より、医薬基盤・健康・栄養研究所(医薬健栄研)において、「創薬ターゲットの枯渇」問題を解決するために、創薬ター
- ゲットがまだ見つかっていない特発性肺線維症(IPF)を対象疾患として、創薬ターゲットを探索する本格的なAI開発に着手した。 • 平成30年7月からは、内閣府による「官民研究開発投資拡大プログラム (PRISM) 」の対象施策に採択され、肺がんを対象疾患に追加し、成果の社会実
- 装に向けて5年間の研究開発目標を設定して文部科学省及び経済産業省が所管する国研等と連携を開始した。
- ◆ 現状
- (データ収集)IPFに関するデータ収集においては、新型コロナの影響により症例収集が半減したため、収集が容易な血液サンプルをオミックス解析の対象と した。今年度中に750症例追加し、累計1,352症例のオミックス解析を実施する予定。 (データ解析)肺がんに関するデータ解析においては、効果的な治療方法が確立されていないパンネガティブ肺がんの発症要因となりうる変異やパスウェイの 変化など創薬ターゲットにつながる多くの新規性の高い知見を得た。
- (結果解釈)英語文献からの情報自動抽出技術を実装したウェブアプリケーションを今年度中にオープンプラットフォームに実装予定である。推論技術開発 は継続中。昨年度見出されたIPFの創薬ターゲット候補のタンパク質について検証を実施中で、有望な創薬ターゲットとなり得ることが確認されている(特許 出願予定)。
- (オープンプラットフォーム)今年度中に国内運用を開始する予定。開始時には今年度はじめまでに完成したAI10種と、IPF(阪大コホート)データを提供 予定である。

研究開発目標

- ① 特発性肺線維症(IPF)及び肺がん患者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する。
- ② 創薬ターゲット/患者層別化バイオマーカーの探索に必要なAI及び知識ベースを開発し、それらの性能を実証する。
- ③ 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。
- ④ 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、疾患データベース、アルゴリズム、開発スキーム等をオープンプラットフォーム化する。

戦略の位置付け等

- AI戦略2021 (令和3年6月11日統合イノベーション戦略推進会議決定)
- II 未来への基盤作り/II-2 研究開発体制の再構築/(1)研究環境整備 具体目標3 世界の研究者から選ばれる、本戦略に即した魅力的な研究開発の制度及びインフラの整備
- III 産業・社会の基盤作り/III-1 社会実装/(1)健康・医療・介護
- 具体目標2 日本が強い医療分野におけるAI技術開発の推進と、医療へのAI活用による医療従事者の負担軽減 具体目標 4 世界最先端の医療AI市場と医療AIハブの形成

資料3-2 「創薬ターゲット探索プラットフォームの構築」の概要②

アドオン 厚生労働省:574百万円、文部科学省:7百万円/元施策 厚生労働省:291百万円

出口戦略

- 本事業で創出される創薬ターゲット候補等の民間企業への導出を図り医薬品の研究開発投資を喚起する。
- ・本事業の実施過程で派生的に作られる技術・成果物を医療システム等に応用するために民間企業への導出を図る。具体的には、電子カルテ関係(医師の入力に係る労務時間を短縮できる電子カルテ入力支援システム、医療テキスト解析システムと疾患知識ベースを用いた症例報告自動作成システム等)や医用画像関連システム(画像アノテーションツール等)への応用を図る。
- この分野では、電子カルテの連携モデルとして、大阪大学附属病院と連携病院との間のシステムを構築できた。また、令和元年に肺がん画像自動アノテーションシステムを企業に導出した他、令和3年には医師が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」を企業と共同開発した。
- 本事業で創出される疾患情報統合データベース、疾患知識ベース及びAIアルゴリズム等の「創薬ターゲット探索のための基盤」について広く産学の研究者に提供する仕組みを構築し、医薬品の研究開発投資を喚起する。
- 本事業により創出される解析アルゴリズムは生活習慣病の先制的介入要因の探索や認知症、精神疾患を始めとする慢性疾患の予後予測など様々なヘルスケア関連のビッグデータ解析に広く活用可能であることから、民間企業が創薬ターゲットの探索等で数多く参加しているLINCやAMED事業等との相互連携もしくは機能統合を図ることによって、広く産業界による利用を促進する。
- 本枠組みを海外にも展開する。

PRISMで推進する理由

- (厚牛労働省)
 - 対象疾患にIPFと深く関連する肺がんを追加し2つに増やすとともに、医薬健栄研が保有するAI開発基盤に加えて理研、産総研などで培ってきた技術を相補 的に活用することにより、元施策事業で目指す事業目標達成の大幅な加速・高度化を期待できる。
- (文部科学省)

理研AIPセンターにて実施するAI等基盤技術の適用による世界初の患者層別化技術の開発に加え、JST-AIPネットワークラボにて実施していた研究課題であるAI技術を用いたマルチオミックス解析基盤技術やバイオマーカー予測法、知識に基づく頑健で高精度な構造的言語処理技術、既承認薬の新しい適応可能疾患の候補を予測する機械学習アルゴリズムなどの研究成果を、アドオンとして創薬ターゲットや患者層別化バイオマーカーの探索に向けたAI開発に発展・展開することで、社会実装や民間企業(製薬企業、診断薬メーカー等)への技術移転等の加速が期待できる。

元施策の加速

- (対象疾患) 難病である特発性肺線維症(IPF)に加え、日本で多くの患者がおり、IPF患者で高頻度に併発する肺がんを加えることができた。
- ・ (症例収集) 肺がんの疾患統合データベースを構築するにあたり、これまで国がんにおいてバイオバンクやがん関連遺伝子のパネル検査等の目的で収集された 肺がんの診療情報及びがん組織を利用し、短期効率的に肺がん疾患統合データベースを構築できた。

また、医療テキストからの情報の抽出にあたりこれまでJSTの事業で培ってきた知識に基づく頑健で高精度な構造的言語処理技術を利用できた。

- (データ解析)理研及びJST事業で培ってきたAI技術を利用して、新規アルゴリズムを構築できた。
- (結果解釈) 産総研で培ってきた自然言語処理及び自動推論技術を利用して以下のことを実施中。
 - ① 英語文献から知識を自動収集するシステムを構築できた。
 - ② 知識ベースから牛体パスウェイ等を自動推論するシステムを構築中。

資料4 令和3年度の成果

○ データ収集

- IPFは、577症例(目標400)の診療情報とオミックス情報(全ゲノム・DNAメチル化解析・RNA-seg)を追加。累計1,352症例(見込み)
- 肺がん症例数は世界最大規模(1,569症例)を達成済み。今年度は欠損値補填としてChIP-seq解析 (50例)、RNA-seq (75例)、全ゲノム解析 (50例)の追加解析実施。

○ データ解析

- IPFについて、既存データベースや論文情報を利用してこれまでに開発してきた創薬ターゲット探索手法4種を統合することにより、より信頼度の高い創薬ターゲット提示システムへと改良した。さらに、新たに開発した予測手法によりIPFとの関連性が示唆される薬剤や遺伝子が追加して見出された。
- 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、症例テキストの時系列検索システムの開発を継続中。医療テキストを自動的に構造化・可視化するシステムを開発し、特許出願中。事業成果を活用した国際コンペティションを開催し、事業成果のアピールに貢献。

○ 結果解釈

- 英語文献からの情報自動抽出技術を実装したウェブアプリケーションのプロトタイプ開発を完了。論文のIDを入力するだけで情報を自動的に抽出し、構造化データとして取得できるようになった。
- ・ 従来の文献解析技術で必要となる学習データを必要とせず、欲しい情報を論文から直接得る新規技術を開発中。
- 情報自動抽出システム及び推論技術開発は継続中。COVID-19関連文献への技術適用も進め、ツール等の公開に向けてウェブサイトを作成。
- 患者層別化AIを用いたIPF臨床情報解析により見出されたIPFの特徴と紐づけられたタンパク質複数個について、肺がん合併IPF患者の肺線維化部位において実際に発現が亢進していることが確認された。
- これまで開発されてきた創薬ターゲット探索手法を用いて既存データベースや論文情報を解析することによって提示されたIPF創薬ターゲット候補は、ヒト家族性IPFの 遺伝子変異を導入した動物実験において有望な創薬ターゲットとなり得ることが確認された。システム及び見出されたIPF創薬ターゲットの両方について特許出願中。
- オミックス解析結果の解釈精緻化を目的に、各種臓器及び血中エクソソームの糖鎖解析及びIPF肺組織シングルセルオミックス解析を実施中。

○ オープンプラットフォーム

- 昨年度にプロトタイプ構築が完了したオープンプラットフォームシステム上で簡便に利用できるようにするため、本事業成果のAIのうち 9 個をアプリ化しシステムに搭載する。 既に搭載済みの患者層別化AIと合わせて合計10個のAIが搭載される予定。
- オープンプラットフォーム上での閲覧・解析を可能にするためには、要配慮個人情報である診療情報に匿名加工が必要となる。阪大より収集したIPF診療情報の匿名加工を完了(見込み)し、プロテオームデータとともにオープンプラットフォームに搭載する。医療画像やテキストなどについても匿名加工準備中。
- オープンプラットフォームの構築に民間企業の意見を反映させるため、これまでの共同研究に加え、IT関連企業1社と製薬企業1社と新たに共同研究を開始した。また、 事業成果をより広く利用して頂くことを目指して、様々な疾患データを収集している他の事業との連携を図り、開発されたAIの適用を拡大していく。また、LINCなど産 学が集うコンソーシアムとも連携し、民間企業にオープンプラットフォーム搭載されたAIを使用してもらい、企業意見を反映させながら改良していく予定。
- オープンプラットフォーム計画の認知度アップのため事業成果の一流国際誌への投稿やマスコミに向けた情報発信。 (令和二年度成果報告会、論文5報、学会発表27件、マスコミ対応4件)

資料5 令和4年度の研究計画

令和4年度目標

- ◆ (データ収集) IP (IPFを含む間質性肺炎):150症例の追加。肺がん:オミックスの追加解析。
- ◆ (データ解析) オープンプラットフォームに搭載するAIの追加開発。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。
- ◆ (結果解釈) 患者由来サンプル及び実験動物を用いて、これまでに見出されてきた創薬ターゲット候補の確からしさを検証する。並びに、文献からの情報自動抽出及び推論技術の完成。
- ◆ (オープンプラットフォーム)AIのアプリ化とオープンプラットフォームへの搭載による一般公開。国内ユーザーの声を反映したシステム構築。

対象施策の研究内容

- ◆ (データ収集)データ収集の継続。最終的目標症例数 IP:1,500症例、肺がん: 1,500症例。
 - ⇒ 令和3年度までのIPの累計症例数 1352症例(見込み)に、IP 150症例分の診療情報及びオミックス情報を追加収集する(累計:1502症例)。
 - ⇒ 肺がんに関しては、令和元年度に目標症例数に達したが、AIによる解析精度の向上のため、pan-negative肺がんを中心にオミックス項目を 追加解析する。
 - ⇒ IPF及び肺がんについて、疾患統合データベースへのデータの収載を完了する。
 - ⇒ 臨床/患者テキストについては、自然言語処理技術の更なる改良を実施する。つまり、これまでに構築したコーパスを拡充し、これを用いた現行 のAIのブラッシュアップにより臨床テキストからの情報自動抽出の精度向上を行う。また、患者テキストからの情報自動抽出や時系列情報の解析を行う技術開発を完了する。
- ◆ (データ解析) 異なる複数の創薬ターゲット探索AIアルゴリズムの適用
 - ⇒ 前年度までに複数のグループが異なるアプローチにより創薬ターゲット探索に有益な基盤技術を開発してきた。令和4年度は対象疾患の臨床情報データベース完成に伴い、これらのAIアルゴリズムをPRISMにて収集した臨床情報全てを対象とした本格的なデータ解析に取り組む。
 - ⇒ 製薬企業などの民間企業にとって魅力的なAIアルゴリズムに作り上げるべく、以下の取り組みを行う。
 - 様々な疾患データを収集している他の事業との連携を図り開発されたAIの適用を拡大する。
 - LINCなど産学が集うコンソーシアムとも連携し、企業ニーズを反映させたAIアルゴリズムの開発を行う。
- ◆ (結果解釈) 文献からの情報自動抽出及び推論技術の完成。
 - ⇒ 患者層別化AI等から洗い出された生体分子を手掛かりにデータウェアハウスのTargetMineや産総研のシステムを用いて疾患パスウェイ等の 更なる分析と実験によるAIの精度評価、創薬ターゲットとしての有望さの評価を行う。
 - ⇒ 英語文献からの情報自動抽出技術によって知識グラフを構築する技術を完成させる。
 - ⇒ 血液又は組織全体のオミックス解析の結果を解釈するために、糖鎖解析を行う。
- ◆ (オープンプラットフォーム) 疾患データ及びAIアプリの追加と海外運用の開始。
 - ⇒ 令和3年度中にオープンプラットフォームに阪大のIPFデータと10種類のAIの搭載を予定しているが、残る肺がん/IPFデータとAIアプリをオープンプラットフォームへ搭載する。また、海外からのデータ及びAIアプリの提供依頼への対応を開始する。

資料6 PRISM実施に伴う事業効果等

〇民間からの貢献額:平成30年度から令和4年度の5年間で約23億円(見込み)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(見込み) 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(見込み) 4.6億円(内訳)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)	
当年度見込み	来年度見込み
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)
 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) in silico創薬に関する共同研究 3件 人件費(67百万円) 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) 	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)	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(8.8百万円)、人件費(5.5百万円)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	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)
(12百万円) (ア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	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)
(50百万円)	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

- ⑧ がん関係の共同研究費 3件 研究費(5百万円)、人件費(90百万 円)
- 計: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) 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が期待できる。(承認薬1剤当たりの研究開発費約 3,000億円)
- 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

人件費(3.8億円)

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。 当年度見込み 来年度見込み • 民間企業からの共同研究等に関する相談・調整等:15社 • 民間企業からの共同研究等に関する相談・調整等:10社

- (外資7社を含む15社の製薬企業等と共同研究についての話し合いを行い、 新たに1社と共同研究契約を締結した。) • AI搭載ITプラットフォームを共同開発し、現在企業と製品化を目指し検証中。
- (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) 8

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)、