

厚生労働科学研究費補助金

政策科学総合研究事業

(臨床研究等 ICT 基盤構築・人工知能実装研究事業)

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

令和元年度 総括研究報告書

研究代表者

水口賢司

国立研究開発法人医薬基盤・健康・栄養研究所

AI 健康・医薬研究センター・センター長

令和 2(2020)年 7 月

目次

1. 統括研究報告書

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-----1

2. 分担研究報告書

- ① 熊ノ郷 淳 大阪大学大学院医学系研究科 呼吸器・免疫内科学-----11
- ② 小倉 高志 神奈川県立循環器呼吸器病センター 呼吸器内科-----15
- ③ 高村 大也 産業技術総合研究所 人工知能研究センター-----17
- ④ 西村紳一郎 北海道大学大学院先端生命科学研究院-----20
- ⑤ 浜本 隆二 国立がん研究センター研究所 がん分子修飾制御額分野-----23
- ⑥ 西岡 安彦 徳島大学大学院医歯薬学研究部 呼吸器・膠原病内科学分野-----25
- ⑦ 奥野 恭史 京都大学医学研究科 人間健康科学系-----28

3. 研究成果の刊行に関する一覧表-----31

厚生労働科学研究費補助金（政策科学総合研究事業）
（臨床研究等 ICT 基盤構築・人工知能実装研究事業）
総括研究報告書

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

研究代表者： 水口賢司（国立研究開発法人医薬基盤・健康・栄養研究所・AI 健康・医薬研究センター・センター長）

要旨の作成

本事業では、「創薬ターゲットの枯渇問題」を克服すべく、動物からではなくヒトの情報から創薬ターゲット分子を探索する AI の開発実装を目的とする。

2019 年度は、創薬標的探索支援データウェアハウス TargetMine に 7 種類の市販/公共データベースおよび知識ベースを統合・拡充した。肺がん手術検体及びバイオプシー検体のオミックス解析を行うとともに、大阪大学病院バイオバンク及び神奈川循環器呼吸器疾患センターで IPF を含む間質性肺炎患者の臨床情報（オミックスデータ及びそれと紐づけられた診療情報）を収集した。miRNA-seq のプロトコルの妥当性評価を行い、適切にデータ収集が可能であることを確認した。昨年度にプロトタイプを開発した患者層別化 AI の改良により、実際に本事業にて収集を進めている臨床情報からデータ駆動的に患者層別化が可能であることを確認した。

参画研究者： 水口賢司（医薬基盤・健康・栄養研究所）、熊ノ郷淳（大阪大学大学院医学系研究科）、小倉高志（神奈川県立循環器呼吸器病センター）、高村大也（産業技術総合研究所）、西村紳一郎（北海道大学大学院先端生命科学研究）、浜本隆二（国立がん研究センター研究所）、西岡安彦（徳島大学大学院医歯薬学研究）、奥野恭史（京都大学医学研究科）

究開発には多額の費用が必要となっており、これが高薬価、ひいては医療費の高騰の要因となっている。

更に、臨床試験段階で期待していた薬効が得られず開発が中断する例が増えていることも問題点として挙げられる。特に医薬品開発の 70～80%が Phase2 で中止となっており、この約 60%が、薬効が得られなかったことが原因との報告がある。つまり、「動物では効くが、ヒトでは効かなかった」という事案が多発している。これは現在の創薬研究開発スキームの限界であると考えられる。

このような現状を打開する解決策として、人工知能（AI; Artificial intelligence）が注目

A. 研究目的

医薬品開発において、近年国内外を問わず創薬ターゲットの枯渇が問題となっている。現在残されているのは高難易度の創薬ターゲットのみであるがために、新薬の研

されている。AIのパフォーマンスと可能性に創薬・医療・ヘルスケア分野が大きな期待を寄せており、今後国際競争が激化することが必至である。

これらの現状を背景に、本事業では、「創薬ターゲットの枯渇問題」を克服すべく、動物からではなくヒトの情報から創薬ターゲット分子を探索する AI の開発実装を目的とする。つまり、臨床情報 (=電子カルテを始めとする診療情報+オミックスデータ) を収集・利用して創薬ターゲットを探索する AI 手法の開発をおこなう。本事業では、対象疾患として難病指定の IPF (特発性肺繊維症) を含む間質性肺炎及び部位別がん死亡者数 1 位である肺癌を選択し、これらの臨床情報収集とそれを支援する基盤構築、異種かつ大量のデータを統合して創薬ターゲット候補となる生体分子群を自動的に抽出する AI 手法の開発を行う。また、本事業で作成される IPF/肺癌の疾患統合データベース、機能分子を特定するための AI 及び知識ベース等を多くの研究者等に利用してもらうための環境 (オープンプラットフォーム) の構築を目指す。

B. 研究方法

1. 既知 (背景) 情報の収集と知識ベース化: 公共・市販のデータベースを TargetMine システムに追加統合した。追加したデータベースは下記の通りである。

- WHO-clinical Trials (臨床データベース)
- HGMD (ヒト遺伝子変異データベース)
- IPF キュレーションベース (定型化された医療情報データベース)
- IPF 情報ベース (論文から抽出した知識ベース)
- PharmaProject (医薬品データベース、臨

床試験情報データベース)

- BioExpress (遺伝子発現情報データベース)
- CTOD (治験情報データベース)

本業務は三菱スペースソフトウェア株式会社への外部委託により実施した。論文からの自動知識抽出を可能にするモデル構築のため、基礎生物学系文献のガイドライン・アノテーション作成および自動キュレーション、および及び臨床系文献のアノテーション基準となるガイドラインの作成を行った。対象となる文献 (IPF・基礎分子系: 121 件、肺癌・基礎分子系 100 件、肺癌・臨床系: 167 件) は当所専門家により選抜され、アノテーション作成は産総研人工知能研究センターのアノテーター 3 名が独立して業務に当たることにより頑健なアノテーションを実施し、随時アノテーション中のデータを解析して一つのエンティティに複数の異なる ID がアサインされていないか確認・修正をおこなった。このプロセスは産総研人工知能研究センターのプロトコルに則って実施された。アノテーションガイドライン作成は当所専門家及び産総研人工知能研究センターの担当研究者・アノテーターと議論の上、本事業の対象疾患である呼吸器疾患の専門用語を適切にアノテーションできるように実践に即して策定した。モデル構築は産総研人工知能研究センターへの委託により実施した。

2. 臨床情報の収集と機械可読表現の開発: 肺癌手術検体及びバイオプシー検体のオミックス解析を行うとともに、大阪大学医学部呼吸器・免疫内科バイオバンク、大阪大学医学部附属病院医療情報部大阪大学病院バイオバンク及び神奈川循環器呼吸

器疾患センターで IPF を含む間質性肺炎患者の臨床情報を収集した。

国立がん研究センターにおいて収集された肺がん手術検体及びバイオプシー検体（検体採取・処理・保存は国立がん研究センターにおけるプロトコルに則って行われた）を用いて、下記のおミックスデータを収集した。おミックスデータ取得はタカラバイオ株式会社への外部委託により実施した。

- ・ 全ゲノム解析
- ・ 全エクソーム解析
- ・ DNA メチル化解析
- ・ ヒストン修飾（ChIP-seq）解析
- ・ トランスクリプトーム（RNA-seq）

大阪大学医学部呼吸器・免疫内科のバイオバンクより、同意取得済みの IPF を含む間質性肺炎患者血清を、また大阪大学医学部附属病院医療情報部より、上記血清にひもつく診療情報（採血日に最も近い日程における診察記録、CT 画像およびその読影所見、血液検査値、呼吸機能検査値およびフローボリュームカーブ、患者基本情報および初診時間診票）を収集した。また、検査の結果、器質的な呼吸器疾患を有しないと診断された方々を健常者とし、同様に血清および診療情報を収集した。なお、血清はバイオバンク番号、診療情報は、研究用の匿名化 ID が付与された形で入手した。

診療情報のうち、診察記録は、必要な情報 120 項目をあらかじめリストして作成したテンプレートを用いて医師が入力もしくはマニュアルでキュレーションを行って構造化データとした。読影所見は、マニュアルあるいは自然言語処理手法を用いて、重要表現にタグ付けを行い、部位と病変のペアからなる構造化データを得た。血液検査値

は、276 項目の重要項目を選択して構造化した。初診時間診票、基本情報は、診察記録のテンプレート項目に情報をキュレーションして追加した。呼吸機能は、匿名化処理の手法構築を行った。

プロテオーム解析

血清 200 μ l から MagCapture isolation kit (Fujifilm Wako)を用いて細胞外小胞を精製した。細胞外小胞中のタンパク質をトリス(2-カルボキシエチル)ホスフィンによる還元・ヨードアセトアミドによるアルキル化を行い、トリプシン消化、脱塩を行った。前処理したサンプルを Data independent acquisition (DIA)法を用いた LC-MS/MS 解析を実施した。データ解析は、DIA 解析ソフト Spectranout を用いて実施し、欠損値には run-wise imputation を行った。Quality control として、15 検体毎に市販血清 1 検体を加えて、サンプル調製からデータ解析までの品質保証を行った。また質量分析の Quality control として HeLa 細胞消化物の DIA 解析を実施した。

神奈川循環器呼吸器疾患センターにおいても、下記の臨床情報（診療情報+おミックスデータまたは肺組織や血液）を収集した。診療情報及び血液は同意を得られた対象疾患患者全員より取得し、採血のタイミングから最も近い診療記録と紐づけた。肺組織は、診断の上でクライオバイオプシーが必要と判断された場合及び外科手術検体が得られた場合においてのみ収集した。クライオバイオプシーを収集する場合、1 患者より数カ所からクライオバイオプシーの採取を行い、縦に半割して一方を病理診断、もう一方をおミックスデータ取得に用いた。採取した肺組織及び血液は、下記のプロト

別紙 3

コルで処理・保存し、1ヶ月に1回当所に輸送（三井倉庫ホールディングス株式会社に委託）し、オミックスデータ取得はタカラバイオ株式会社に委託した。

診療情報

- ・ 診療記録 (97 項目)
- ・ 所見
 - 超音波 (5 項目)
 - 気管支鏡 (19 項目)
 - 外科生検 (16 項目)
 - CT 画像検査 (19 項目)
- ・ 検査
 - 血液+尿 (103 項目)
 - 生理機能 (35 項目)

CT 画像

肺組織

- ・ クライオバイオプシー:採取後チューブに回収し、速やかに-80℃で保存した。
- ・ 外科手術検体 (VATS):採取後チューブに回収し、速やかに-80℃で保存した。
- ・ 上記と紐づけられた病理所見
- ・ 上記と紐づけられた病理画像からの特徴量

血液

- ・ 血清 (プロテオーム解析用):採血後、室温で静かに5回転倒混和し、室温で30分間静置した後にスイング式ローターを用いて、室温、1,300×gで10分遠心した。上清をチューブに250μLずつ分注して速やかに-80℃で保存した。
- ・ 血漿 (miRNA トランスクリプトーム解析用):採血後、室温で静かに10回転倒混和し、スイング式ローターを用いて、室温、1,300×gで10分遠心した。上清をチューブに250μLずつ分注して速やかに-80℃で保存した。
- ・ 末梢血単核細胞 (PBMC、全ゲノム解析

用):採血後、室温で静かに5回転倒混和し、スイング式ローターを用いて室温、1,600×gで20分遠心した。まとめて採血する場合、最初の検体から最後の検体の時間が1時間以内になるようにした。遠心後、ゲルバリアの上にある細胞層を乱さないようにして血漿層を約半分吸引し、パスツールピペットを用いて、細胞層全量をチューブに回収して速やかに-80℃で保存した。

肺組織からの核酸 (DNA・RNA) 同時抽出
肺組織サンプル (56 症例、80 検体) について、DNA・RNA 同時抽出キット (NucleoSpin RNA、NucleoSpin RNA/DNA Buffer Set) を用いて核酸抽出を行った。抽出された核酸はバイオアナライザーを用いて RNA 分画における DNA 混入の有無を確認し、RNA 分画は NucleoSpin RNA/DNA Buffer Set のマニュアル通りに DNase 処理を行なった。

肺組織からのマルチオミックスデータ収集

1. 全ゲノム解析 (depth 30)

(1)TruSeq DNA ライブラリの作成

アコースティックソルビライザー Covaris(コバリス社)を用いて物理的に数百bpに断片化したDNAの両末端を平滑化とリン酸化処理をした後、3'-dA 突出末端処理を行い、Index 付きアダプターを連結した。その後、アガロースゲル電気泳動によりアダプターを連結したDNAをサイズ選別し、それをを鋳型とし、PCRによる増幅を行い、シーケンスライブラリとした。ライブラリの品質はAgilent2100 BioAnalyzerを用いて測定した。

(2)シーケンス解析

Illumina社のNovaSeq6000を用いて150base長ペアエンドシーケンスを行い、シーケン

サー付属ソフト (NovaSeq Control Software v1.6.、Real Time Analysis (RTA) v3.4.、bcl2fastq2 v2.20) を用いて塩基配列 (リード配列) を得た。タグ配列に基づき塩基配列(リード配列)を検体ごとに分類して fastq ファイルを取得した。

2. DNA メチル化解析

Illumina 社の Infinium MethylationEPIC BeadChip を用いて、Infinium HD Methylation Protocol Guide, Manual Protocol (15019519 v01) に従い、ゲノム DNA を用いてメチル化解析を実施した。

(1) Bisulfite 処理

ゲノム DNA (250ng) を EZ DNA Methylation Kit(ZYMO RESEARCH 社)を用いて Bisulfite 変換し、精製、回収を行った。

(2)全ゲノム増幅、断片化、及び精製

Bisulfite 処理した DNA はアルカリ処理後に酵素による全ゲノム増幅を行なった。その後、酵素による断片化、イソプロパノール沈殿による精製の後、バッファーに再懸濁した。これを熱変性させ、Infinium MethylationEPIC BeadChip にアプライして、48°C で約 23 時間ハイブリダイゼーションを行なった。

(3)一塩基伸長反応とスキャニング

ハイブリダイゼーション後、BeadChip をバッファーで洗浄し、一塩基伸長反応によってプローブ末端に一塩基の標識ヌクレオチドを導入、ハイブリダイズした DNA を変性・除去し、取り込ませた標識ヌクレオチドに対する蛍光色素標識抗体染色を行い、蛍光イメージを取得した。

(4)データ解析

取得した蛍光イメージデータを GenomeStudio/Methylation Module を用いて、Background Subtraction 及び Normalizaion を

実施して解析を行なった。

3.トランスクリプトーム (RNA-seq)

(1) TruSeq Stranded mRNA ライブラリの作製

検体より PolyA+RNA を単離し、断片化により得られる RNA を鋳型とした逆転写反応により一本鎖 cDNA を合成した。これを鋳型として、dUTP を取り込ませた二本鎖 cDNA を合成した。得られた二本鎖 cDNA の両末端を平滑化・リン酸化処理した後、3'-dA 突出処理を行い、Index 付きアダプターを連結した。アダプターを連結した二本鎖 cDNA を鋳型とし、dUTP を持つ鎖を選択的に増幅しないポリメラーゼにより PCR 増幅を行い、シーケンスライブラリとした。ライブラリの品質は Agilent2100 BioAnalyzer を用いて測定した。

(2)シーケンス解析

Illumina 社の NovaSeq6000 を用いて 150base 長ペアエンドシーケンスを行い、シーケンサー付属ソフト (NovaSeq Control Software v1.6.、Real Time Analysis (RTA) v3.4.、bcl2fastq2 v2.20) を用いて塩基配列 (リード配列) を得た。タグ配列に基づき塩基配列(リード配列)を検体ごとに分類して fastq ファイルを取得した。

miRNA-seq のフィージビリティースタディ ー (FS)

神奈川循環器呼吸器疾患センターにおいて収集された血漿 24 症例から 250µl ずつプールし、FS 用サンプルを調製した。これを容量(100 または 200µl)、凍結融解の回数(1,2,3 回)、スパイク (QIAseq miRNA Library QC Spike-Ins : RNA 抽出前に添加)の有無について検討した。RNA 抽出は QIAGEN miRNeasy micro kit を用いた。QIAseq miRNA

Library QC Spike-Ins (nuclease free water 500µl で懸濁したもの) と QIAzol を混合し mixture を作成したのちに、各サンプルに 700µl を添加して RNA 抽出後、14µl の nuclease free water で溶出した。ライブラリ調製キットは QIAseq miRNA Library Kit を使用し、input 量は 5µl とした。こうして調製したライブラリの quality check はバイオアナライザーと qPCR(2nM サンプルを使用)にておこなった。miRNA-seq は Nextseq 75bp シングルエンドで行い、1 サンプルあたり 2000 万リード以上取得することとした。RNA 抽出から miRNA-seq までの一連の作業は株式会社 DNA チップ研究所に委託した。

収集された臨床情報は、当所で構築を進めている疾患統合データベースに順次格納する。当該データベースは 2018 年度に MongoDB (オープンソースソフトウェアのドキュメント指向データベース) を用いて構築した (三井情報株式会社への委託)。2019 年度は、当該データベースに格納されたデータの統計量 (データの特徴を要約した数値) を表示する機能を実装した (ライフマティックス株式会社への委託)。

3. 統合データ解析の手法開発とそれを用いた新規創薬ターゲット候補の同定：昨年度にプロトタイプを開発した患者層別化 AI の改良を実施した。本事業で収集を進めている臨床情報は、ある項目が当てはまるか当てはまらないかを 0 (当てはまらない) または 1 (当てはまる) で表現した数字が記載されているものと実測値 (例：SpO2) が記載されているものが混在する。2020 年度は、これを入力データとして受け入れることが可能となるように改良した。プログ

ラムは python で実装した。大阪大学病院にて収集された IPF を含む間質性肺炎患者 363 症例の診療情報 (日常診療の電子カルテ:110 項目、血液検査:276 項目、CT 画像読影所見:3342 項目。欠損値は下記の方法で補填した。①項目が当てはまるか否かを 1 または 0 で表現している場合、欠損値には 0 を補填、②実測値が記載される場合、欠損値には k-nearest neighbors で補填) 及びそれを紐づけられた血清エクソソーム プロテオームデータ (1941 種類のタンパク質) を用い、改良した患者層別化 AI の動作確認をおこなった。

(倫理面への配慮)

本研究は、医薬基盤・健康・栄養研究所ならびに分担研究機関において倫理審査、承認を得た後、ヒトゲノム・遺伝子解析研究に関する倫理指針及び、人を対象とする医学系研究に関する倫理指針に従って遂行した。

C. 研究結果

1. 既知 (背景) 情報の収集と知識ベース化：計画通りに公共・市販データベースの統合を完了した。論文から分子・病態に関する知識抽出を行う AI 開発のために必要な言語リソースの構築において、産総研人工知能研究センターと共同で肺がん・基礎分子系要旨データ 100 件のエンティティ (disorder、cell、pharmacological substance) アノテーションはほぼ終了し、一貫性の確認・修正作業を実施している。IPF・基礎分子系要旨データに対して、生命現象イベント (gene expression、conversion、pathway、regulation) に関するアノテーション用ガイドラインを作成し、IPF・基礎分子系要旨データ 121 件に対してエンティティ (GGPs、

analytical entity、subject、cell、cell component、有機化合物、無機化合物) のアノテーションとノーマリゼーション ID における一貫性の確認を行った。イベントについても、gene expression など比較的シンプルなものに対してアノテーションを実施した。肺がん・臨床系論文 167 件のアノテーションについては、アノテーションガイドライン作成後に再開する計画へと変更し、現在中断している。言語リソース構築に関する研究成果全般の詳細については、研究分担者(高村大也)による報告書を参照されたい。

2. 臨床情報の収集と機械可読表現の開発：肺がん手術検体及びバイオプシー検体のオミックス解析の研究結果全般の詳細については、研究分担者(浜本隆二)による報告書を参照されたい。

大阪大学医学部呼吸器・免疫内科のバイオバンクおよび大阪大学医学部附属病院医療情報部より、IPF を含む間質性肺炎および器質的な呼吸器疾患がないと診断された健常者について、下記の臨床情報を取得した。

- ・ 血清 602 症例
- ・ 診療記録+基本情報 594 症例
- ・ 血液検査値 594 症例
- ・ 読影所見 538 症例
- ・ 初診時間診票 141 症例
- ・ 画像 323 症例

血清は、エクソソームを分離したのち、エクソソーム内プロテオームデータを取得し、診療情報は、方法に記載した手法を用いてそれぞれ構造化した。初診時間診票、基本情報は、診察記録のテンプレート項目に情報をキュレーションした。呼吸機能データについては、匿名化処理の手法構築のみを行った。

これらの臨床情報を用い、患者層別化 AI での解析を試行した(本研究結果は次項「3. 統合データ解析の手法開発とそれを用いた新規創薬ターゲット候補の同定」を参照)。大阪大学バイオバンクの臨床情報収集全般の詳細については、研究分担者(熊ノ郷淳)による報告書を参照されたい。

神奈川循環器呼吸器病センターより、下記の臨床情報を取得した(2020年2月20日現在)。

- ・ 研究同意取得 452 名
- ・ 血液 370 名
- ・ 肺組織(クライオバイオプシー・外科手術検体) 82 検体
- ・ 臨床診断 310 名

肺組織 56 症例(80 検体)を対象に下記のマルチオミックスデータを取得した。

- ・ 全ゲノム解析(depth 30)
- ・ DNA メチル化解析
- ・ トランスクリプトーム(RNA-seq)

神奈川循環器呼吸器病センターの臨床情報収集全般の詳細については、研究分担者(小倉高志)による報告書を参照されたい。

肺組織からの核酸(DNA・RNA)同時抽出
神奈川循環器呼吸器病センターにおいて収集された肺組織 56 症例(80 検体)について DNA・RNA 抽出を行ない、シーケン스에必要量を確保することができるか否かを検討した。

miRNA-seq のフィージビリティスタディー(FS)

患者 24 人分の血漿をプールして調製したサンプル(容量 2 点(100 μ L, 200 μ L)、凍結融解 3 回、スパイクの有無について各 3 サンプル用意し、合計 36 サンプル)について

B.研究方法に記載した方法で miRNA-seq を実施した。その結果、シーケンスの Phred score (リード中の各ポジションにおける塩基のシーケンス QC スコア) が 75bp の全領域において 30 (シーケンシングが良好に行われたことを示す基準値) を上回ることを確認した。2020 年 6 月現在、miRBase (miRNA のデータベース) に登録されているヒト miRNA の数は 2500 強であり、本研究では全てのサンプルにおいて 2000 以上の miRNA が検出された。総リードに含まれる miRNA の割合は、容量 100 μ l のサンプルでは 15-20%、容量 200 μ l のサンプルでは 20-30%であった。凍結融解の結果からは、1 回目と 2 回目の間で多くの miRNA が分解されていることがわかった。更に、サンプルの凍結融解による miRNA 発現量への影響をより詳細に調査するため、凍結融解による分解を受けることがすでに報告されている miR-1 (Glinge, C. et al (2017). Stability of circulating blood-based microRNAs - pre-analytic methodological considerations. PloS one, 12(2), e0167969.) の発現量を各サンプル間で比較した。その結果、容量 100 μ l・200 μ l 両方において miR-1 の発現量が減少する傾向が認められた。特に、容量 200 μ l の場合においては、凍結融解 1-2 回では各実験条件における miR-1 発現量の平均値が 150 前後であったのに対して凍結融解 3 回では平均値が 100 前後であり、1 回目と 3 回目の間では p -value < 0.05 (Welch's t test) で有意差があった。続いて、凍結融解の回数間で増減した miRNA の数を volcano plot により確認した。その結果、スパイクなしのサンプルでは、凍結融解数の違いにより有意に発現量が異なる miRNA が一定量検出されたのに対して、スパイクありでは発現量に有意差のある miRNA の数が減

少し、凍結融解による影響が認められにくい傾向があった。

3. 統合データ解析の手法開発とそれを用いた新規創薬ターゲット候補の同定：大阪大学医学部付属病院にて収集された IPF を含む間質性肺炎患者 363 症例の診療情報及びそれを紐づけられた血清エクソソームプロテオームデータを改良した患者層別化 AI に供したところ、「診療情報項目と紐づけられたタンパク質群」を出力し設計通りの動作を確認した。

D. 考察

1. 既知 (背景) 情報の収集と知識ベース化：産総研人工知能研究センターより、「現時点のアノテーション・ツール (brat) では、一つのエンティティに複数の ID を付与できない。しかしながら、実際のアノテーションでは、一つのエンティティに複数の ID を付与しないといけないケースが多々あるので、どのように解決するか、アノテーター、システム開発の担当者とも議論する必要がある。」「生命現象イベントは非常に複雑なので、文献データをどのようにテキスト・アノテーションしていくかが課題となる。複雑なアノテーションは、アノテーターによるアノテーションのコストも高くなり、時間を費やすだけで、ネットワークなどの関係抽出をする際にも精度が落ちる可能性があるため、アノテーションしやすいスキームを設計する必要がある。そのためにも、肺疾患の専門家、テキストマイニングの専門家の意見を取り入れることが望まれる。」と現状における課題が提示されており、当所肺疾患の専門家の意見を反映させた言語リソース構築を継続する必要がある。

2. 臨床情報の収集と機械可読表現の開発
大阪大学病院バイオバンクで収集した臨床情報は計画内容をほぼ完了し、2020年度は患者層別化 AI による本解析を主軸としてデータ解析による知識抽出に注力する段階に到達していると言える。神奈川循環器呼吸器病センターにおける臨床情報収集では、特に間質性肺炎の中でもどのような症例を重点的に収集するべきかを随時収集状況に関する情報共有をしながら柔軟に対応する必要がある。

miRNA-seq のフィージビリティスタディー (FS)

miRNA はいずれのサンプルにおいても2000種類を超える miRNA を検出することができた。総リード数に対する miRNA の割合は容量100µlよりも200µlの方が高く平均26%であった。また、スパイクの有無では差がなかった。凍結融解を繰り返すと検出される miRNA の割合や特定の miRNA の発現量が減少することを確認した。また、スパイクの有無によりグローバルな miRNA 発現プロファイルに影響が出る可能性が示唆された。

スパイクの有無によって凍結融解の回数による影響の見え方が異なる点については、スパイクを入れることによって相対的に miRNA が占める割合が低下し、各サンプル間での差異が結果的に小さくなることによると考えられた。

3. 統合データ解析の手法開発とそれを用いた新規創薬ターゲット候補の同定
大阪大学バイオバンクの全 602 症例データを用いた本解析により得られる出力については、本事業で拡充した TargetMine を用いたより多様な情報源を基にした情報検索・

分析により結果解釈を実施する他、研究分担者（熊ノ郷淳）のグループに属する呼吸器内科の医師らによる結果解釈を取り入れることにより創薬標的探索に有益な情報を得られるものと期待される。

E. 結論

1. 既知（背景）情報の収集と知識ベース化：TargetMine の拡充が順調に進んだと判断し、2020年度は公共データベース2点（PoSSuM、MGeND）の追加統合のみの実施とする。産総研人工知能研究センターとの委託研究について、2020年度に計画内容を完了するべく、言語リソース構築についても完了・モデル構築において学習データとして利用可能な状態で産総研人工知能研究センターに提供することを目標とする。

2. 臨床情報の収集と機械可読表現の開発
国立がん研究センターにおける肺がんの手術検体及びバイオプシー検体を用いたマルチオミックス解析は、全エクソーム解析やトランスクリプトーム解析（RNA-seq）において世界最大規模のデータベース構築を達成したことから、2020年度は欠損データの取得を実施することとする。大阪大学病院バイオバンクにおける IPF を含む間質性肺炎の臨床情報収集は、2020年度に全 602 症例分の血清中エクソソームのプロテオームデータ及びそれと紐づけられた診療情報の収集をもって完了する。2019年度に神奈川循環器呼吸器病センターにて収集された肺組織を用いたマルチオミックスデータは、2020年度にプロセッシング・解析を実施するほか、引き続き肺組織・血液・診療情報の収集を継続する。

miRNA-seq のフィージビリティスタディー

一 (FS)

血漿中 miRNA の網羅的測定にあたり、血液サンプルの収集プロトコルや RNA 抽出から miRNA-seq に至るプロトコルは適切に設定されていると判断した。

本研究結果を基に、

- ・ 容量：200µl
- ・ 凍結融解は1回に留める
- ・ スパイクは使用しない

という方針でプロトコルを設定することとする。miRNA-seq に至る全行程のプロトコルの妥当性を確認することができたと判断し、2020 年度において血漿中 miRNA の測定を実施する計画を具体化させることとした。

3. 統合データ解析の手法開発とそれを用いた新規創薬ターゲット候補の同定：2019 年度に実施した患者層別化 AI の改良により、本事業にて収集を進めている臨床情報からデータ駆動的に患者層別化ルールを出力できることを確認した。これにより、臨床情報の収集・構造化が完了次第、当該 AI を用いた本解析を実施することとした。

F. 健康危機情報

なし

G. 研究発表

1. 論文発表

1. 伊藤眞里, 長尾知生子, 藤原大, 水口賢司, AI 創薬に向けた医療ビッグデータベースの統合と解析, 月刊ファームステージ, Vol.19, No11 2 月刊 21-24, 2020
2. 夏目やよい, 水口賢司, 新薬創出を加速する AI の開発, **Precision Medicine** Vol.3 No.5, 2020 (印刷中)

2. 学会発表

1. 水口賢司, Artificial intelligence-based drug discovery: challenges and applications to target identification and pharmacokinetic modelling, 第4回トランスオミクスシンポジウム, 徳島, 2019.11.14
2. 水口賢司, 疾患・創薬研究におけるデータベースと AI 活用, 日本臨床試験学会 第11回学術集会総会, 東京, 2020.2.14
3. 水口賢司, 創薬研究への AI 活用, **ILSI Japan 先端技術シンポジウム**, 東京, 2020.2.21
4. 夏目やよい, Artificial intelligence for patient stratification based on clinical information, **The 1st International Symposium on Human InformatiX**, 京都, 2020.2.28
5. T. Fujiwara, T. Ogawa, H. Nakagawa, H. Hirata, I. Nagatomo, M. Itoh, Y. Takeda, A. Kumanogo, K. Mizuguchi, Multimodal feed-forward neural network coupled with convolutional neural network for severity diagnosis of idiopathic interstitial pneumonia using medical information and computed tomography images, **The 1st International Symposium on Human InformatiX**, 京都, 2020.2.28

H. 知的財産権の出願・登録状況

1. 特許取得
該当事項なし
2. 実用新案登録
該当事項なし
3. その他
特記事項なし

2019年度研究実施報告書

(対象期間 2019年11月13日～2020年3月31日)

担当課題：肺がん統合データベース構築及びAI技術を用いたオミックス解析

研究機関名：国立研究開発法人国立がん研究センター

研究責任者：浜本 隆二 rhamamot@ncc.go.jp

【1. 実施内容】

実施項目1

1-1 研究目的

動物ではなく、ヒトの情報から創薬ターゲット分子を探索するAIの開発実装を行うことを目的とする。対象疾患としては、特に肺がんに焦点を当てて行う。

1-2 研究概要・要旨

肺がんに対する臨床情報（電子カルテを始めとする診療情報＋患者検体のオミックスデータ）を収集・利用して創薬ターゲットを探索するAI手法の開発を行う。

1-3 実施内容

- ① オミックスデータの取得：国がんで行っているバイオバンク事業等と連携し、それらの事業で取得された肺がん検体からDNA/RNAを取得し、医薬健栄研に提供した。また、肺がん検体のChIP-seq解析を行った。
- ② 肺がん患者の臨床情報のデータベース化：①の肺がん検体と紐付いた臨床情報を収集し、AIの入力データとして利用するためにクレンジング等の作業を行った。
- ③ 統合データ解析のAI手法開発：①のオミックスデータとそれに紐付く②の臨床情報を用い、肺がんの本態解明、ひいては創薬ターゲット候補を見出すためのAI手法の開発を行う。

1-4 結果、成果等

- ① データ収集に関しては、全エクソーム解析及び診療情報に関しては521症例新たに肺がんデータベースに追加し、総計1569症例となり世界最大規模のデータベースとなった。その他にも全ゲノム解析は242症例（総計272症例）、DNAメチル化解析は160症例（総計430症例）、ChIP-seq解析は100症例（総計100症例）、トランスクリプトーム（RNA-seq）解析は654症例（総計1494症例）が肺がんデータベースに追加された。
- ② 呼吸器内科・呼吸器外科にある電子カルテ情報を、ゲノム・エピゲノム情報と紐づけた統合データベースを構築した。
- ③ 肺がんオミックス解析を目的とした新規アルゴリズムを4種類開発し、学会発表及び論文発表を行った。

【2. 外部発表、論文投稿等】

[学会発表]

1. 浜本 隆二. がんの統合的解明を目指したオミックス情報の階層的ネットワークに対する機械学習・深層学習技術の応用: 第39回医療情報学連合大会 (招待講演)
2. 浜本 隆二. 人工知能 (AI) 技術の医療応用への取組-医用画像解析を中心として-: 第37回日本脳腫瘍学会学術集会 (招待講演)
3. 小林 和馬, 三宅 基隆, 渡辺 裕一, 浜本 隆二. 臨床のリアル・ワールド・データを利用した人工知能技術の開発を加速するための統合的なデータ・プラットフォームの構築: 第2回日本メディカル AI 学会学術集会
4. 浅田 健, 小林 和馬, 高橋 慧, 高澤 建, 小松 正明, 金子 修三, 浜本 隆二. 肺がん予後予測を施行したマルチオミックス解析: 第2回日本メディカル AI 学会学術集会
5. 高澤 建, 浜本 隆二. 知識行列を用いたスパース正則化付き NMF による DNA メチル化データ解析手法の開発: 第2回日本メディカル AI 学会学術集会
6. Syuzo Kaneko, Ryuji Hamamoto. Toward the application of precision medicine: Multi-omics analysis including clinical ChIP-seq datasets may reveal pivotal regulation of gene expression in cancer: 第2回日本メディカル AI 学会学術集会

[論文投稿]

1. Kobayashi K, Murakami N, Takahashi K, Inaba K, Igaki H, Hamamoto R, Itami J. A Population-based Statistical Model for Investigating Heterogeneous Intraprostatic Sensitivity to Radiation Toxicity After ¹²⁵I Seed Implantation. *In Vivo*, 33:2103-2111 (2019)
2. Asada K, Kobayashi K, Joutard S, Tubaki M, Takahashi S, Takasawa K, Komatsu M, Kaneko S, Sese J, Hamamoto R. Uncovering Prognosis-related Genes and Pathways by Multi-omics Analysis in Lung Cancer. *Biomolecules*, 10:524 (2020)
3. Hamamoto R, Komatsu M, Takasawa K, Asada K, Kaneko S. Epigenetics Analysis and Integrated Analysis of Multiomics Data, Including Epigenetic Data, Using Artificial Intelligence in the Era of Precision Medicine. *Biomolecules*, 10:62 (2020)
4. Isago, H, Mitani A, Mikami Y, Horie M, Urushiyama H, Hamamoto R, Terasaki Y, Nagase T. Epithelial Expression of YAP and TAZ Is Sequentially Required in Lung Development. *Am J Respir Cell Mol Biol*, 62 :256-266 (2020)

【3. 知財化について】

該当なし

2019年度研究実施報告書

(対象期間 2019年11月13日～2020年3月31日)

担当課題 : 「新薬創出を加速する人工知能の開発」
機関名 : 神奈川県立循環器呼吸器病センター
研究責任者 : 小倉 高志

【1. 実施内容】

1 研究目的

特発性肺線維症(IPF)及び間質性肺炎(疑いを含む)疾患に関する診療情報および患者サンプルの収集を行い、臨床情報(診療情報および患者サンプルから得たオミックスデータ)を統合する医療情報統合データベースを完成させること。

2 研究概要・要旨

実臨床の情報により検証し、創薬ターゲット分子の発見・同定へと導く。対象疾患としては、指定難病の一つである特発性肺線維症を主体とした間質性肺炎と肺がんとし、それらに対する創薬ターゲット分子を提案する。

3 実施内容 IPFを含む間質性肺炎(疑いを含む)の診断となった患者から臨床情報、各種検査結果の抽出、研究用に血液10mLを採取する。また、研究同意を得られた患者からの気管支鏡検査、外科的肺生検からの肺組織採取を行う。臨床情報は、初診時間診票の記載データと電子カルテ日常診療医師入力記録から抽出を行っている。各種検査結果は、胸部CT、X線画像、血液検査データ(血算、肝機能、腎機能、炎症マーカーなど)、呼吸機能検査および生理機能検査データ、画像読影所見(CT画像、X線画像に付随の読影所見)、病理所見から得られた結果を抽出した。外科生検とクライオ生検による得られた肺組織からDNA、RNAの同時抽出を施行

DNA抽出に関して56症例80検体で施行し、問題なく抽出可能であった。

RNA抽出に関して56症例80検体で施行し、問題なく抽出可能であった。

研究開始から2020年3月31日まで493人の研究同意取得を行っており、別紙(表1)のように血液数425人、生検数103人、診断に関しては、別紙(表2)のように352人の患者が臨床診断を得ている。未診断例も日々、診断を行ってきている。今後も順次、気管支鏡検査、外科的肺生検を施行する予定としているが、ここまでの採取検体の解析に関しては、特にコロナウイルス感染症の影響は受けていないと考えている。

4 結果、成果等

成果：間質性肺炎は診断にある程度の情報量と検討時間が必要であり、同意取得者がどのような診断になるかは、その後の検査結果と診断合議によるところが大きい。本来、無治療の初診患者に対して情報抽出、血液、組織検体採取を施行する予定であったが、すでに治療薬（抗線維化剤、ステロイド、免疫抑制剤）が導入されている患者群も入っている。

同意取得時点で診断はできないためにかなり診断区分が多種にわたっている（表2参照）。IPFは全体の13%程度となっているが、今後さらにMDD診断を行うことで症例数は増えると考えられる。

費用：検体収集のための遠心分離機、-80℃での保存が可能な冷凍庫等、また、実施に必要な運営要員の雇用等が主な使途である。

【2. 外部発表、論文投稿等】

該当なし。

【3. 知財化について】

該当なし

表1 当院における間質性肺炎の同意書数・採血数・生検数

	同意書数	採血数	生検数
2019/11月	84	9	24
2019/12月	231	105	25
2020/01月	95	141	25
2020/02月	49	92	14
2020/03月	34	78	15
2019年度小計	493	425	103

表2 当院でのMDD診断

01_IPF	65
02_UCIP	121
03_cHP	58
04_NSIP	45
05_others	63
06_未解析	141
総計	493

以上

2019年度研究実施報告書

(対象期間 2019年11月13日～2020年3月31日)

担当課題：新薬創出を加速する人工知能の開発

研究機関名：国立大学法人大阪大学

研究責任者：熊ノ郷 淳 kumanogo@imed3.med.osaka-u.ac.jp

【1. 実施内容】

実施項目 1

1-1 研究目的

課題のプロジェクトは、PRISM(官民研究開発投資拡大プログラム)のサイバー空間基盤技術に採択されたプロジェクトとして、国立研究開発法人 医薬基盤・健康・栄養研究所(医薬健栄研)が中心となり、省庁連携にて推進するプログラムである。その計画は、大量の疾患情報を整理、分類、蓄積した AI を構築し、臨床情報(診療情報及びマルチオミックスデータ)から分子と疾患の因果関係の推定や創薬ターゲット探索につながる仮説創出を行う。

1-2 研究概要・要旨

ヒト由来培養細胞・動物を用いた実験及び実臨床の情報により検証し、創薬ターゲット分子の発見・同定へと導く。対象疾患としては、指定難病の一つである特発性肺線維症と肺がんとし、それらに対する創薬ターゲット分子を提案する。

1-3 実施内容

当院および連携各病院においては、間質性肺炎(特発性肺線維症を含む)患者から同意を得て、①検査画像、②呼吸機能検査値およびフローボリュームカーブ(FVC)、③血液検査値、④画像読影所見、⑤診療記録、⑥初診時間診票、⑦血清サンプルを採取し医薬健栄研に提供する。その際、医療情報部が構築する CDGS システムと連携する。

1-4 結果、成果等

#1 電子カルテ入力テンプレートの作成と実装

成果：2017年度は、共同研究開始と同時にカルテテンプレートの作成に取り掛かり、わずか2か月間にて初診問診票とカルテテンプレート作成を完成した。PRISMにおいて欠かせないパーツで

ある患者情報の収集システム構築を整備した。今年度は、近畿中央呼吸器センター、羽曳野医療センター、さらに刀根山医療センターへ OCRnet を拡大し、関連病院のサンプルを収集できるインフラ整備を国内で先駆けて整備することに成功した。

費用（消耗品）：特に本目的での使用はありません。

#2 臨床検体の採取

成果：2017年度に呼吸器・免疫内科で継続する血清バンクシステムを構築し、2か月間で合計80検体の血清保存、そのうち肺線維症は5検体採取した。血清からエクソソームの単離について、種々のエクソソーム単離法の比較、エクソソームの質や量の比較をすることで今後の大量オミクスのための予備解析を進めた。具体的には、エクソソームの単離方法は、超遠心法だけでなく、EV-second やマグキャプチャー法などが報告されている。そこで、当科では予備検討として、どの方法が適切であるか？さらにどのようにしてエクソソームを検証するのかを確認した。プロテオミクスの結果から、マグキャプチャーによる単離法が、最も純度が高いことを見出した。さらに、検証法として、Nanosight や qNano、電子顕微鏡にてエクソソームを確認できること、さらにエクソソームの同定に必要な各種マーカー抗体による WB から、同定法も確認した。以上より、本研究において、今後収集する貴重な患者検体におけるエクソソームの単離法と検証法を確認できた。当院にて収集した血清検体を基盤研に提供した。当該サンプルは、基盤研にて単離法含めて構築した次世代プロテオミクスや他のオミクス（マルチオミクス）にて解析予定であり、AI のためのビッグデータとなる。

費用：エクソソームの単離&血清分離 (1)、エクソソームの確認に必要な試薬 (2) を購入した。大量サンプルを同時に検証するために、EV-Second L70 など単離回収に必要な試薬に充当した。

#3 解析における助言

成果：本研究目的は、カルテやエクソソームのマルチオミクスを活用して、病態解明、創薬、バイオマーカー開発に役立てることである。ところが、本研究にてフォーカスするエクソソームが、炎症性肺疾患の肺病変を反映するのか？そもそも、そのような物質がエクソソームに含まれることすら不明である。そこで、肺線維症の患者からエクソソームを単離して、プロテオミクスが肺の炎症や線維化病態を反映することを検証した。ただし、当科にて構築した血清バンクは極めて貴重であり、エクソソーム単離や予備検討に使用することは困難である。そこで、当科が今までに構築したマウス肺線維症モデル（ブレオマイシン誘導モデル）を、ヒトにおける検証の予備段階として解析することで、マウス肺内病変がエクソソームの蛋白に反映されること、さらに血液中のエクソソームの蛋白群が、肺線維化病態（TGF-beta signaling）を示唆することも見出した。さらに、発見したエクソソームの BM 候補蛋白は、肺線維化病変において発現が増加するだけでなく、有意な相関を見出した。

本成果を受けて、2018年度はマウスモデルで得られた内容を、ヒトにおいて検証へと進めた。基盤研が開発した次世代プロテオミクス DIA によるヒト血清エクソソームと肺線維症組織は解析を終えて、エクソソームの内容が肺組織を反映することが示唆された。

以上、我々が得た知見は、PRISMにおいて大量解析するサンプルとしてエクソソームを活用する合理性となるデータを提供するものであり、解析における助言として十分である。しかし、エクソソームの肺線維症における病的意義は不明であることから、単離エクソソームが上皮細胞や炎症細胞に与える役割を検討する必要がある。そこで、in vitro の解析としてマウスや患者エクソソームを、肺構成細胞への投与実験を行った。具体的には、マクロファージ・血管内皮細胞 (HUVEC)、上皮細胞にエクソソームを投与し、細胞増殖 (Cell Counting Kit-8)、細胞死 (APC Annexin V)、サイトカイン産生を解析した。

費用：血清エクソソームのプロテオミクス (受託解析) やマウスモデルにおける病理解析 (受託解析) 以外に、エクソソーム投与による細胞への影響を検討するために用いる in vitro 解析 (3) や細胞培養に必要な種々の消耗品 (4)、さらにはヒト実験の予備検討としてのマウス実験 (5) 成果報告のための学会参加 (6) が主な使用用途である。

【2. 外部発表、論文投稿等】

Dysregulated Expression of the Nuclear Exosome Targeting Complex Component Rbm7 in Nonhematopoietic Cells Licenses the Development of Fibrosis.

Fukushima K, Satoh T, Sugihara F, Sato Y, Okamoto T, Mitsui Y, Yoshio S, Li S, Nojima S, Motooka D, Nakamura S, Kida H, Standley DM, Morii E, Kanto T, Yanagita M, Matsuura Y, Nagasawa T, Kumanogoh A, Akira S.

Immunity. 2020 Mar 17;52(3):542-556.e13. doi: 10.1016/j.immuni.2020.02.007.

Pathological and therapeutic implications of eosinophil-derived semaphorin 4D in eosinophilic chronic rhinosinusitis.

Tsuda T, Nishide M, Maeda Y, Hayama Y, Koyama S, Nojima S, Takamatsu H, Okuzaki D, Morita T, Nakatani T, Kato Y, Nakanishi Y, Futami Y, Suga Y, Naito Y, Konaka H, Satoh S, Naito M, Izumi M, Obata S, Nakatani A, Shikina T, Takeda K, Hayama M, Inohara H, Kumanogoh A.

J Allergy Clin Immunol. 2020 Mar;145(3):843-854.e4.

LRRK2 regulates endoplasmic reticulum-mitochondrial tethering through the PERK-mediated ubiquitination pathway.

Toyofuku T, Okamoto Y, Ishikawa T, Sasawatari S, Kumanogoh A.

EMBO J. 2020 Jan 15;39(2):e100875.

Metagenome-wide association study of gut microbiome revealed novel aetiology of rheumatoid arthritis in the Japanese population.

Kishikawa T, Maeda Y, Nii T, Motooka D, Matsumoto Y, Matsushita M, Matsuoka H, Yoshimura M, Kawada S, Teshigawara S, Oguro E, Okita Y, Kawamoto K, Higa S, Hirano T, Narazaki M, Ogata A, Saeki Y, Nakamura S, Inohara H, Kumanogoh A, Takeda K, Okada Y.
Ann Rheum Dis. 2020 Jan;79(1):103-111.

【3. 知財化について】

該当なし

2019年度研究実施報告書

(対象期間 2019年11月13日～2020年3月31日)

担当課題 : ヒト組織由来エクソソームの糖鎖プロファイリングによる新しいバイオマーカーの探索

研究機関名 : 国立大学法人北海道大学

研究責任者 : 西村紳一郎 shin@sci.hokudai.ac.jp

【1. 実施内容】

実施項目1 ヒト臓器由来エクソソーム糖鎖データベース構築に向けた基盤技術の確立

1-1 研究目的

ヒト組織由来エクソソーム糖鎖データベースを構築するため、エクソソーム作製のための基本的な方法とそれらの糖鎖プロファイリング技術を確立する。本基盤技術とデータベースを活用して鋭敏な診断用バイオマーカーも有効な治療法も存在しない特発性肺線維症などの難病を含む多くの疾患における新しい診断技術と治療法を開発することを主たる目的とする。

1-2 研究概要・要旨

臓器・組織の糖鎖プロファイリングならびにそれらの培養細胞からのエクソソーム作製法とその糖鎖構造プロファイリング法を確立した。

1-3 実施内容、結果、成果等

健康な（正常）ヒト臓器由来培養細胞からのエクソソーム作製法とその糖鎖構造プロファイリング法（細胞由来エクソソーム糖鎖解析プロトコル）を確立するため、マウスおよびヒトがん培養細胞を用いてエクソソーム調製プロトコルを種々検討した。

その結果、ヒト培養細胞由来エクソソームの糖鎖構造を解析・同定するにあたり、その基本的な構造情報として（1）細胞全体の糖鎖プロファイル（whole *N*-glycan profile）、（2）細胞膜表面糖鎖のプロファイル（surface *N*-glycan profile）、さらに、（3）実際に単離精製したエクソソームの糖鎖（exosomal *N*-glycan profile）を確保する必要があること、また、（4）エクソソームを放出・エンリッチする方法や培地などの諸条件を統一する必要があることなどを明らかにした。

実施項目2 特発性肺線維症のバイオマーカー探索

2-1 研究目的

肺線維化病態マウス組織検体（全20検体）で前年度初歩的な解析で確認した糖鎖構造を確定して指標とするバイオマーカー探索研究をさらに進展させる。

1-2 研究概要・要旨

正常マウスおよび医薬基盤研チームにより作製された肺線維化病態マウス臓器・組織試料を用いて糖鎖構造解析を行い疾患特異的な糖鎖構造を探索した。

1-3 実施内容、結果、成果等

肺線維化病態マウス肺組織サンプル（医薬基盤研チームより供与された特発性肺線維症病態マウス肺組織凍結粉碎試料、DBA/2Crslc および MRL-lpr/lpr : BLM 投与群とコントロールを含めて全 20 検体）すべてについての糖鎖プロファイリングを完了した。

その結果、マウスにおける病態組織（細胞）全体の糖鎖プロファイル (*N*-glycan profiles from whole lung tissue) を初めて獲得した。IPF 疾患マウスに特異的な糖鎖群における発現レベルの顕著な上昇を確認した（実施項目 1 の研究成果と合わせて論文投稿に向けて準備中）。

実施項目 3 マウスの主要臓器由来エクソソーム糖鎖データベース構築に向けた基盤技術の確立

3-1 研究目的

正常マウス臓器・組織由来培養細胞からのエクソソーム作製法とその糖鎖構造プロファイリング法（マウス細胞由来エクソソーム糖鎖解析プロトコル）を確立する。この新たなプロトコルによってマウスの主要臓器・組織由来エクソソームの糖鎖構造のプロファイリングを開始する。正常なマウス臓器・組織についての網羅的エクソソーム糖鎖情報が獲得できれば、マウス血液中のエクソソーム糖鎖プロファイルとの比較から由来臓器・組織を推測する実証試験を実施する。「研究項目 2：特発性肺線維症のバイオマーカー探索」への応用を視野に研究を進める。

3-2 研究概要・要旨

マウス培養細胞由来エクソソームの糖鎖解析プロトコルを確立した。マウスの主要臓器・組織由来エクソソームの糖鎖構造のプロファイリングを開始した。

3-3 実施内容、結果、成果等

実施項目 1 で得られた「ヒト培養がん細胞由来エクソソームの糖鎖構造解析」の結果を踏まえて、マウス肺がん細胞（3LL）をモデルとして、（1）エクソソームを調製する際の培地血清由来エクソソーム（FBS exosomes）の影響を評価すると同時に、（2）細胞全体の糖鎖プロファイル（whole *N*-glycan profile）と実際に単離精製したエクソソームの糖鎖プロファイル（exosomal *N*-glycan profile）の違いを比較した結果、ヒト培養がん細胞で観察されたと同様の傾向があることを確認した。また、（3）超遠心法に代わるエクソソームの簡便・迅速エンリッチ法について比較検討し、それぞれの処方におけるメリット・デメリットを明らかにした。さらに、（4）正常マウスの主要臓器・組織・細胞の糖鎖解析を開始した（2020 年 2 月）。エクソソームの標準的な糖鎖解析法に関する最初の論文投稿に向けて準備中。

【2. 外部発表、論文投稿等】

本研究課題に関連する論文を含む。

1. Soma O., Hatakeyama S., Yoneyama T., Saito M., Sasaki H., Tobisawa Y., Noro D., Suzuki Y.,

Tanaka M., Nishimura S-I., Harada H., Ishida H., Tanabe K., Satoh S., Ohyama C., “Serum *N*-glycan profiling can predict biopsy-proven graft rejection after living kidney transplantation” *Clin. Exp. Nephrol.* **24**, 174-184 (2020)

【3. 知財化について】

特になし。

2019年度研究実施報告書

(対象期間 2019年11月13日～2020年3月31日)

担当課題 : 新薬創出を加速する人工知能の開発

研究機関名 : 国立研究開発法人産業技術総合研究所

研究責任者 : 高村 大也 takamura.hiroya@aist.go.jp

【1. 実施内容】

実施項目 1 (言語リソース構築)

1-1 研究目的

文献からの分子、病態の知識抽出を行うために、AIに学習させるための言語リソースを構築する。

1-2 研究概要・要旨

文献からの分子や病態に関する知識、特にプロテイン名や疾患名などのエンティティとそれらの間の関係を自動的に抽出するために、抽出を行うAIを学習するための訓練データとしてアノテーション付きコーパスを構築する。コーパス構築は、アノテーションガイドラインの作成および、実際のアノテーションを含む。

1-3 実施内容

文献としてはIPFに関する基礎分子系の論文を選択し、そのアノテーションガイドラインを作成する。また、作成したアノテーションガイドラインに基づき、文献要旨に対してアノテーションを行う。

1-4 結果、成果等

IPF基礎分子系ガイドライン作成に関しては、Regulation(制御)などの複雑な生命現象イベントについて、テキストマイニングの専門家や肺疾患の専門家にアノテーション例を複数提案して、意見を取りまとめながら、アノテーション用ガイドラインの作成を継続した。

IPFの基礎分子系要旨データのアノテーションに関しては、121件の要旨データに対して、Gene expression(遺伝子発現)など比較的シンプルなイベントのアノテーションを開始した。また、アノテーション用の文献要旨データ30件を新たに選抜し、アノテーション用のデータセットとして準備した。この際に、主に Genes and gene products(GGPs; 遺伝子や蛋白質など遺伝子産物)に関する記述の多い要旨を選抜した。

実施項目 2 (文献情報と生物学情報の融合)

2-1 研究目的

新規標的候補の同定に向けた文献情報と生物学情報との融合を目指し、当研究所で開発されたデータベース等を用いて検索・推論基盤を構築する。

2-2 研究概要・要旨

文献から抽出された情報と生物学情報との融合のために、抽出情報との関連探索技術を開発し、またデータベースの開発・更新・維持を行う。

2-3 実施内容

タンパク質の基質結合(候補)部位に関するデータベース PoSSuM に関し、検索・推論基盤の構築を行う。また、肺疾患に関し、学術論文や既存データベースなどからタンパク質や相互作用に関する情報を抽出する。

2-4 結果、成果等

当研究所で開発されたタンパク質の基質結合(候補)部位に関するデータベース PoSSuM を例として、検索・推論基盤の構築に向けた準備を開始(PoSSuM の更新に着手)している。データの劇的な増加に対処するため導入した新規計算機を用い、更新作業を行っている。これと並行し、ケーススタディとして、対象疾患の発症や治療に関わるとされる免疫システムについて、特に肺線維症に着目して学術論文や既存データベースなどから関連するタンパク質や相互作用等に関する情報を抽出した。またデータベース化に向けデータ編集方法を検討した。

実施項目3 (文献からのエンティティおよび関係の抽出)

3-1 研究目的

基礎生物科学系文献、および臨床系文献から、蛋白質、疾患などのエンティティ、およびそれらの間の関係を自動的に抽出する技術を開発する。

3-2 研究概要・要旨

文献から、蛋白質、疾患などのエンティティを抽出する技術、エンティティをデータベースにリンクする技術、エンティティ間の関係を推定する技術を開発する。また、文献のキュレーションデータを直接推定する技術の開発も行う。

3-3 実施内容

研究期間内では、エンティティリンク技術の開発を進め、項目1で開発されたコーパスおよび、一般公開されている別ドメインのコーパスで性能評価を行う。また、文献からのキュレーションデータの推定については、一部の属性について、実際にプロトタイプツールを構築する。

3-4 結果、成果等

エンティティリンクについて、モデル構築、実験などをさらに進めた。開発手法は、mention detection, candidate generation, candidate ranking の三段階から成る。このうち、candidate generation の段階で、高精度に候補を挙げられるように手法を改良した。また、開発した手法は、性能評価実験において良好な結果が得られている。文献からのキュレーションデータの獲得は、一部の属性については自動的に高精度に抽出できるようになった。また、簡易的なツールという形に実装した。

【2. 外部発表、論文投稿等】

該当なし

【3. 知財化について】

該当なし

2019年度研究実施報告書

(対象期間 2020年1月1日～2020年3月31日)

担当課題：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

研究機関名：国立大学法人 京都大学

研究責任者：教授 奥野 恭史

【1. 実施内容】

実施項目 1

1-1 研究目的

既存公開症例データベースを人工知能技術に適用可能な状態に整備し、そのデータを用いて主に深層学習技術の一つである Graph Convolutional Network (GCN) 及び説明可能人工知能技術の一つであるベイジアンネットワーク (BN) などを利用・拡張・改良することにより創薬ターゲット探索・推定が行えるアルゴリズムを開発する。

1-2 研究概要・要旨

既存公開がん症例データベースである The Cancer Genome Atlas (TCGA)、Genomics of Drug Sensitivity in Cancer (GDSC)、及び Cancer Cell Line Encyclopedia (CCLE) に登録されている各種がん症例及びがん細胞株薬剤応答データを、人工知能・機械学習アルゴリズムに適応可能な状態に整備する。

1-3 実施内容

まず既存公開がん症例データベースである The Cancer Genome Atlas (TCGA)、Genomics of Drug Sensitivity in Cancer (GDSC)、及び Cancer Cell Line Encyclopedia (CCLE) に登録されている各種がん症例及びがん細胞株薬剤応答データを、人工知能・機械学習アルゴリズムに適応可能な状態に整備する。

次に、そのデータを用い、主に GCN 及び BN を用いて創薬ターゲット因子の予測のためのアルゴリズムの改良を行う。GCN は大量のデータからグラフ（ネットワーク）で表現される関係性の予測に使用できるもので、すでに遺伝子ごとのゲノム変異情報、薬剤、及びタンパク間相互作用情報から構築したネットワークを用いた関係性予測で一定の予測能力を確認している。これを創薬ターゲットとなる因子の予測に改良を行う。BN は因子間の因果関係をグラフとして明に表す教師なしの機械学習アルゴリズムである。そのため疾患や薬剤の作用機序解明に有用である。データから推定されたグラフ構造から創薬ターゲットとなりうる因子の特徴を見出すことにより、創薬ターゲット予測へと応用可能であると考え、そのグラフ上の特徴量の創出を行う。

1-4 結果、成果等

効率的に薬剤反応性の予測やバイオマーカーの推定を行うためには、ゲノムやオミクスデータ、薬剤（化合物）、臨床情報の情報を組み合わせた超高次元のデータを扱う必要がある。そのため、適切に **feature engineering**、次元圧縮、特徴選択等の技術を用いて、本質的なデータを抽出することが重要である。本年度は、データを組み合わせる方法として、複数種類のデータを入力するマルチモーダルアプローチと複数の出力を同時に扱うマルチタスクアプローチの二つのプロトタイプ実装を行い、公共利用可能なデータ（CCLE/GDSC 等）を用いて評価することのできる環境を整えた。

上述データを用いた GCN を用いて創薬ターゲット因子の予測のためのアルゴリズムの改良として、本年度は、文献や既存の公共データベースにある網羅的な情報（文献空間の情報）を実験によって得られた実測空間の情報を統合した機械学習手法のプロトタイプ開発を行った。また、次年度以降にこれらを容易に利用できるようオープンソースの GCN フレームワークである **kGCN** (<https://github.com/clinfo/kGCN>) の基盤開発を行った。

BN を用いた創薬ターゲット因子の特徴量創出については、これまで肺がん培養細胞株を用いたトランスクリプトーム・ネットワーク解析により TGF β 投与によって引き起こされる上皮間葉移行 (Epithelial-Mesenchymal Transition: EMT) の重要サブネットワークの抽出に成功している。この EMT サブネットワークを TCGA 肺がん患者に適応することにより、患者予後の層別化が可能なことまではわかっている (Tanaka et al., 2020)。この BN による EMT サブネットワークを用いて、今回新たに患者毎のサブネットワークの可視化を試み、それに成功した (図 1)。前述の論文 (Tanaka et al., 2020) では EMT サブネットワークより算出した患者サンプルの枝貢献量行列をクラスタリングすることにより患者の層別化を行ったが、この図はこのクラスタリング結果から抽出した各クラスターの代表的な 4 人の患者の EMT サブネットワークを新しく開発した方法で可視化したものである。この図で示した患者毎のネットワークの違いは TGF β 投与に反応するサブネ

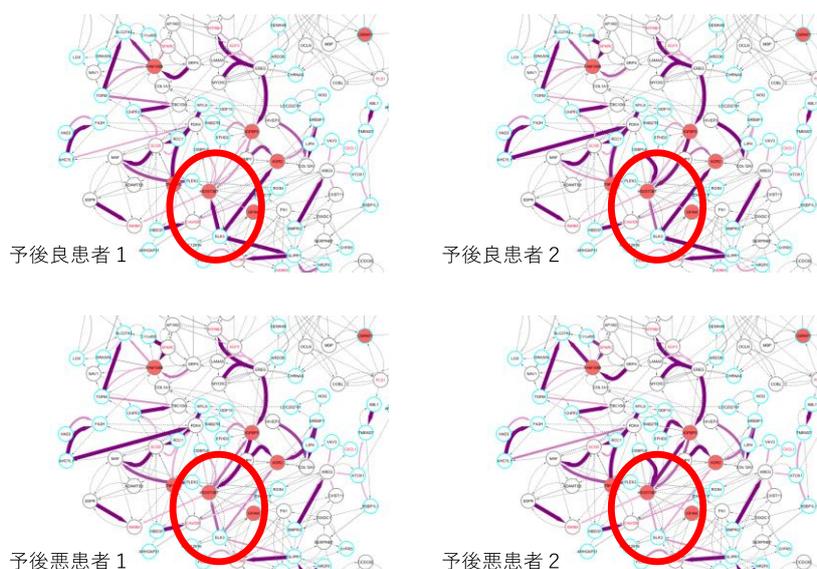


図 1. 患者別 EMT サブネットワークの部分拡大図。赤丸で示した周辺が特に変化が大きいことがわかる。

ネットワークにおける各患者の違いを示しており、特定のハブ遺伝子周辺に変化が大きいことが観察される。これは推定したネットワーク構造の評価だけでは抽出不可能で、したがって、これらの特徴的な遺伝子が創薬ターゲットになりうると考えられる。またこれによって可視化されたサブネットワークの違いに基づき患者毎に異なる創薬ターゲットを考えることや薬剤感受性の違いの説明も可能になると考えられる。今後はこれらの特徴量の薬剤ターゲットとしての可能性を解析評価予定である。

【2. 外部発表、論文投稿等】

Tanaka, Y., Tamada, Y., Ikeguchi, M., Yamashita, F., Okuno, Y., System-based differential gene network analysis for characterizing a sample-specific subnetwork, *Biomolecules*, 10(2), 306, 2020.

【3. 知財化について】

奥野恭史、玉田嘉紀. 特願 2020-002923 「特徴ネットワーク抽出装置、コンピュータプログラム、特徴ネットワーク抽出方法及びベイジアンネットワーク分析方法」 (出願日：2020/01/10)

2019年度研究実施報告書

(対象期間 2020年1月1日～2020年3月31日)

担当課題：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
研究機関名：国立大学法人徳島大学
研究責任者：西岡 安彦

【1. 実施内容】

実施項目 1

1-1 研究目的

間質性肺炎合併肺癌特有の間質細胞プロファイルを見出し、その肺癌の病態との関りを検討すると同時にデータベース構築を行い AI を用いた新規創薬ターゲット探索の礎を作る。

1-2 計画概要

間質性肺炎合併および非合併肺癌患者の肺癌切除検体を用いて、①肺の健常部位、②肺癌部位、③間質性肺炎部位より得られた間質細胞のシングルセル解析を実施して遺伝子発現を網羅的に検討することで(図1)、間質性肺炎合併肺癌に特徴的な間質細胞プロファイルを明らかにするとともに、そのオミックス解析の結果からがんの進展に影響を与える可能性がある生体分子のデータベース構築を行う。

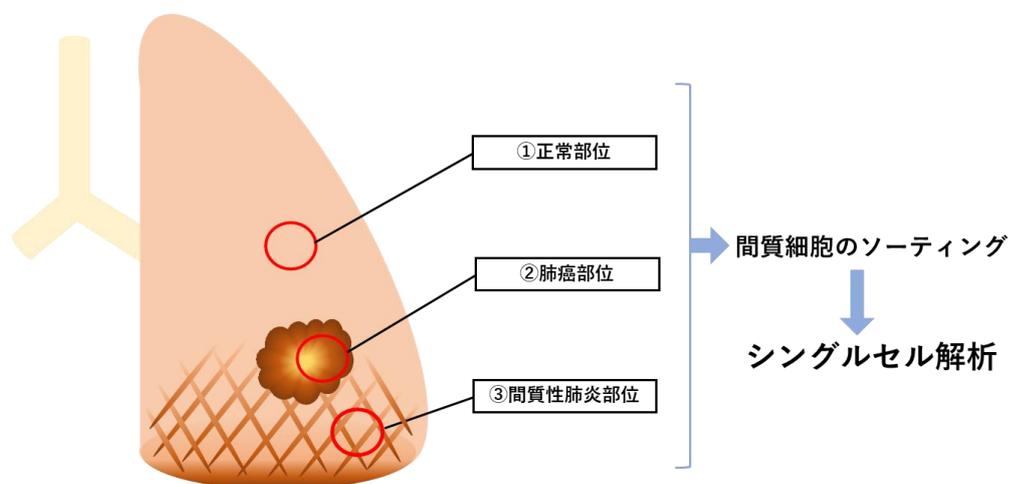


図1. 本検討の概要

1-3 実施内容

1-2で計画したヒト臨床研究を行うため研究計画書を作成し、2020年1月に徳島大学病院医学系研究倫理審査委員会に提出した。2020年2月17日に承認され（受付番号：3647）、ただちに研究を開始した。

シングルセル解析、とくに sc-RNA seq および sc-ATAC seq の技術を習得するため、マウス肺を検体として解析を行い、肺組織からの単細胞浮遊液の作成からはじまり、10x Genomics 社 Chromium システムを用いたライブラリー作成、シーケンス、Cell Ranger 解析までの一連の実験手法を検討した。この際にシングルセル解析の手法で、均一な emulsion が作成できない、サーマルサイクラーの使用上のトラブルがあり PCR ができない、作成したライブラリーの DNA 量に変動がみられるといった問題点があがり、それぞれ、実験中の温度管理やピペットチップの選択、サーマルサイクラーの使用機種を選択、死細胞の効率的な除去で解決できる可能性が考えられた。

またヒト検体においてシングルセル解析に向けた実験方法を検討した。肺がん手術検体より得られた組織を利用し、解析に必要な組織サイズを把握した。また線維芽細胞を含む間質細胞は既報告にあるように Epcam⁺CD45⁻CD31⁻細胞として採取することとした（Cell Rep. 2018; 22: 3625-40.）。肺がんの外科的切除検体を使用して Epcam⁺CD45⁻CD31⁻細胞をソーティングするための gating strategy を作成し、解析に必要な細胞数を Cell Sorter を用いて収集できることが分かった。

本年度のトラブルシューティングをしたシングルセル解析の手法の習熟により、次年度に向けて実際の肺がん手術検体でシングルセル解析を行う準備ができた。

1-4 結果、成果等

本年度は肺がん検体のシングルセル解析に向けた手法の確認を行った。また肺がん検体から間質細胞を Epcam⁺CD45⁻CD31⁻細胞として採取する gating strategy を確認した（図2）。これらにより次年度は具体的に肺がん検体のシングルセル解析を行い、間質性肺炎合併肺がん特有の間質細胞のプロファイル特定やデータベース構築を行い AI を用いた新規創薬ターゲット探索の礎を作る。

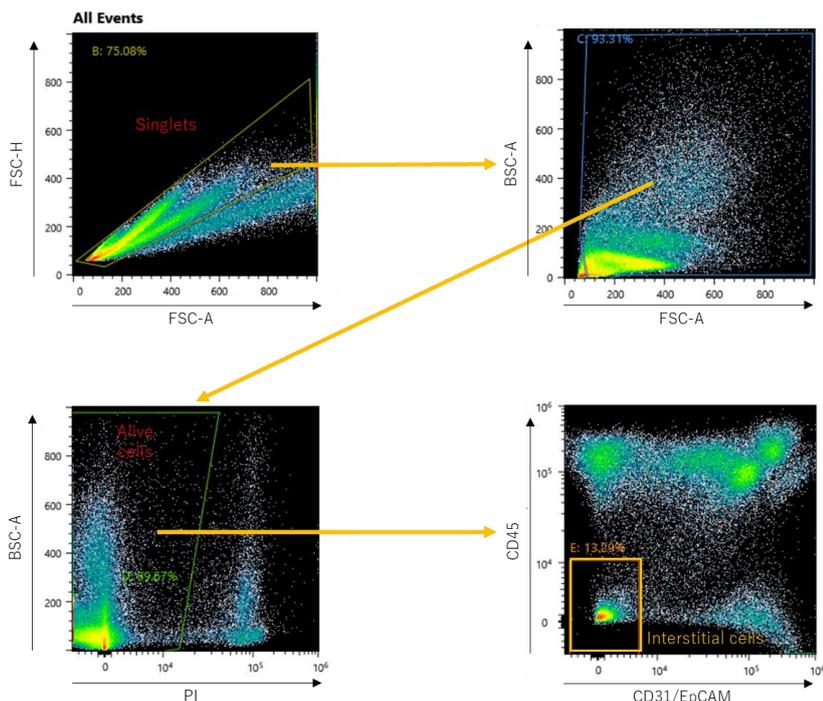


図2. ヒト肺サンプルからの間質細胞ソーティングのための Gating strategy

【2. 外部発表、論文投稿等】

本年度の外部発表および論文発表はない。

【3. 知財化について】

現時点で知財化に該当する成果はない。

III. 研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
伊藤眞里, 長尾知生子, 藤原大, 水口賢司	AI創薬に向けた医療ビッグデータベースの統合と解析	月刊 ファーマステージ	Vol.19, No.1	21-24	2020
夏目やよい, 水口賢司	新薬創出を加速するAIの開発	Precision Medicine	Vol.3 No.5	印刷中	2020
Fukushima K, Satoh T, Sugihara F, Sato Y, Okamoto T, Mitsui Y, Yoshio S, Li S, Nonaka S, Motooka D, Nakamura S, Kida H, Standley DM, Morii E, Kaneko T, Yanagita M, Matsuura Y, Nagasawa T, Kumanogoh A, Akira S	<u>Dysregulated Expression of the Nuclear Exosome Targeting Complex Component Rbm7 in Nonhematopoietic Cells Licenses the Development of Fibrosis.</u>	Immunity	52(3)	542-556.e13	2020
Tsuda T, Nishide M, Maeda Y, Hayama Y, Koyama S, Nojima S, Takamatsu H, Okuzaki D, Morita T, Nakatani T, Kato Y, Nakanishi Y, Futami Y, Sugaya Y, Naito Y, Konaka H, Satoh S, Naito M, Izumi M, Obata S, Nakatani A, Shikina T, Takeda K, Hayama M, Inohara H, Kumanogoh A	<u>Pathological and therapeutic implications of eosinophil-derived semaphorin 4D in eosinophilic chronic rhinosinusitis.</u>	J Allergy Clin Immunol	145(3)	843-854.e4	2020
Toyofuku T, Okamoto Y, Ishikawa T, Sasawatari S, Kumanogoh A.	<u>LRRK2 regulates endoplasmic reticulum-mitochondrial tethering through the PERK-mediated ubiquitination pathway.</u>	EMBO J	39(2)	e100875	2020

Kishikawa T, Maeda Y, Nii T, Motooka D, Matsumoto Y, Matsushita M, Matsuoka H, Yoshimura M, Kawada S, Teshigawara S, Oguro E, Okita Y, Kawamoto K, Higa S, Hirano T, Narazaki M, Ogata A, Sasaki Y, Nakamura S, Inohara H, Kumano A, Takekida K, Okada Y.	Metagenome-wide association study of gut microbiome revealed novel aetiology of rheumatoid arthritis in the Japanese population.	Ann Rheum Dis	79(1)	103-111	2020
Soma O., Hatakeyama S., Yoneyama T., Saito M., Sasaki H., Tobisawa Y., Nishimura S-I., Harada H., Ishida H., Tanabe K., Satoh S., Ohyama C	Serum N-glycan profiling can predict bioprophyl. sy-proven graft rejection after living kidney transplantation	Clin. Exp. Nephrol.	24	174-184	2020
Kobayashi K, Murakami N, Takahashi K, Inaba K, Igaki H, Hamamoto R, Itami J	A Population-based Statistical Model for Investigating Heterogeneous Intraprostatic Sensitivity to Radiation Toxicity After ¹²⁵ I Seed Implantation. In Vivo	Biomolecules	10	524	2019
Asada K, Kobayashi K, Joutard S, Tubaki M, Takahashi S, Takasawa K, Komatsu M, Kaneko S, Sese J, Hamamoto R	Uncovering Prognosis-related Genes and Pathways by Multi-omics Analysis in Lung Cancer. Biomolecules	Biomolecules	10	524	2020

Hamamoto R, Komatsu M, Takasawa K, Asada K, Kaneko S.	Epigenetics Analysis and Integrated Analysis of Multiomics Data, Including Epigenetic Data, Using Artificial Intelligence in the Era of Precision Medicine	Biomolecules	10	62	2020
Isago, H, Mitani A, Mikami Y, Horie M, Urushiyama H, Hamamoto R, Terasaki Y, Nagase T	Epithelial Expression of YAP and TAZ is Sequentially Required in Lung Development	Am J Respir Cell Mol Biol,	62	256-266	2020
Tanaka, Y., Tamada, Y., Ikeguchi, M., Yamashita, F., Okuno, Y.	System-based differential gene network analysis for characterizing a sample-specific subnetwork,	Biomolecules,	10(2)	306	2020

特許

発明者氏名	発明の名称	出願番号
小林和馬、三宅基隆、浜本隆二	アノテーション支援装置、アノテーション支援方法及びアノテーション支援プログラム	特願2019-010287
奥野恭史、玉田嘉紀	特徴ネットワーク抽出装置、コンピュータプログラム、特徴ネットワーク抽出方法及びベイジアンネットワーク分析方法	特願2020-002923

厚生労働大臣 殿

令和2年 4 月 9 日

機関名 国立研究開発法人
医薬基盤・健康・栄養研究所

所属研究機関長 職 名 理事長

氏 名 米田 悦啓



次の職員の令和元年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 政策科学総合研究事業（臨床研究等 ICT 基盤構築・人工知能実装研究事業）

2. 研究課題名 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

3. 研究者名（所属部局・職名） AI 健康・医薬研究センター・センター長

（氏名・フリガナ） 水口 賢司（ミズグチ ケンジ）

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	医薬基盤・健康・栄養研究所	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針（※3）	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	医薬基盤・健康・栄養研究所	<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること （指針の名称：)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査に場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合は委託先機関：)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由：)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> （有の場合はその内容：)

（留意事項） ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

厚生労働大臣 殿

機関名 国立大学法人大阪大学

所属研究機関長 職名 大学院医学系研究科長

氏名 森井 英一 印



次の職員の令和元年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 政策科学総合研究事業（臨床研究等 ICT 基盤構築・人工知能実装研究事業）
2. 研究課題名 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
3. 研究者名（所属部局・職名） 大学院医学系研究科・教授
 （氏名・フリガナ） 熊ノ郷 淳 ・クマノゴウ アツシ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入（※1）		
	有	無	審査済み	審査した機関	未審査（※2）
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針（※3）	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	大阪大学医学部附属病院	<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること （指針の名称： ）	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

（※1）当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他（特記事項）

（※2）未審査の場合は、その理由を記載すること。

（※3）廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由： ）
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合は委託先機関： ）
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> （無の場合はその理由： ）
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> （有の場合はその内容： ）

（留意事項） ・該当する□にチェックを入れること。
 ・分担研究者の所属する機関の長も作成すること。

令和2年5月8日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 横浜市金沢区富岡東6丁目16番1号
地方独立行政法人
所属研究機関長 職名 神奈川県立病院機構
神奈川県立循環器呼吸器病センター
氏名 所長 田尻道彦印



次の職員の令和元年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 政策科学総合研究事業 (臨床研究等ICT基盤構築・人工知能実装研究事業)
2. 研究課題名 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(19AC5001)
3. 研究者名 (所属部局・職名) 呼吸器内科・部長
(氏名・フリガナ) 小倉 高志・オグラ タカシ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	神奈川県立循環器呼吸器病センター研究倫理審査委員会	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	神奈川県立循環器呼吸器病センター研究倫理審査委員会	<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和2年 5月 11日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 国立研究開発法人産業技術総合研究所

所属研究機関長 職名 理事長

氏名 石村 和彦



次の職員の令和元年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 1. 研究事業名 政策科学総合研究事業（臨床研究等 ICT 基盤構築・人工知能実装研究事業）
- 2. 研究課題名 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
- 3. 研究者名 (所属部局・職名) 人工知能研究センター・知能情報研究チーム 研究チーム長
(氏名・フリガナ) 高村 大也・タカムラ ヒロヤ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。
(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

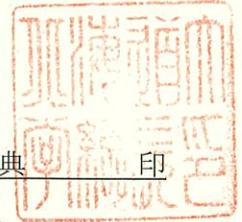
(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

厚生労働大臣
 (国立医薬品食品衛生研究所長) 殿
 (国立保健医療科学院長)

機関名 北海道大学

所属研究機関長 職名 総長職務代理

氏名 笠原 正典



次の職員の令和元年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 政策科学総合研究事業 (臨床研究等 ICT 基盤構築・人工知能実装研究事業)
2. 研究課題名 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
3. 研究者名 (所属部局・職名) 大学院先端生命科学研究院・教授
 (氏名・フリガナ) 西村 紳一郎・ニシムラ シンイチロウ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
 ・分担研究者の所属する機関の長も作成すること。

2020年4月1日

厚生労働大臣 殿

機関名 国立研究開発法人国立がん研究センター

所属研究機関長 職名 理事長

氏名 中釜 斉 印



次の職員の令和元年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 1. 研究事業名 政策科学総合研究事業
- 2. 研究課題名 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
- 3. 研究者名 (所属部局・職名) 研究所 がん分子修飾制御学分野・分野長
(氏名・フリガナ) 浜本 隆二・ハマモト リュウジ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	国立がん研究センター	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	国立がん研究センター	<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

令和2年 4月 1日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 徳島大学

所属研究機関長 職名 学長

氏名 野地 澄晴



次の職員の令和元年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 1. 研究事業名 政策科学総合研究事業（臨床研究等 ICT 基盤構築・人工知能実装研究事業）
- 2. 研究課題名 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
- 3. 研究者名 (所属部局・職名) 大学院医歯薬学研究部・教授
(氏名・フリガナ) 西岡 安彦・ニシオカ ヤスヒコ

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	徳島大学	<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

2020年4月8日

厚生労働大臣
(国立医薬品食品衛生研究所長) 殿
(国立保健医療科学院長)

機関名 国立大学法人京都大学

所属研究機関長 職名 京都大学大学院医学研究科長

氏名 岩井 一宏 印



次の職員の平成 年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 政策科学総合研究事業（臨床研究等 ICT 基盤構築・人工知能実装研究事業）
2. 研究課題名 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
3. 研究者名 （所属部局・職名）国立大学法人京都大学 医学研究科 人間健康科学系専攻 教授
（氏名・フリガナ）奥野 恭史

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
ヒトゲノム・遺伝子解析研究に関する倫理指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
人を対象とする医学系研究に関する倫理指針 (※3)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」や「臨床研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。