

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）

大規模データの利活用研究の加速のための研究

令和4年度 総括・分担研究年度終了報告書

研究代表者 木村 映善

令和4年（2022年）年 5月

目 次

I. 総括研究年度終了報告	
大規模データの利活用研究の加速のための研究	----- 1
木村 映善	
II. 分担研究年度終了報告	
1. 大規模データの利活用研究の加速のための研究	----- 3
渋谷 哲朗	
2. 千年カルテの匿名加工医療情報を利用した入院後敗血症の予測モデルの開発における 課題と対策の検討	----- 4
松田 敦義	
3. 大規模データの利活用研究の加速のための研究	----- 8
佐々木 香織	
4. 千年カルテの匿名加工医療情報を利用した多医療機関の電子カルテに対するシーケン ス解析に関する研究	----- 11
横田 治夫	
III. 研究成果の刊行に関する一覧表	----- 13

令和4年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)
総括研究年度終了報告書

大規模データの利活用研究の加速のための研究

研究代表者 木村 映善 (愛媛大学医学部 教授)

研究要旨

認定事業者を利用した研究を加速する施策の提言につなげることをめざし、認定事業者の潜在的な利用者に関する意識調査、海外の医療情報の収集・活用に関する調査、匿名加工医療情報を用いたAI研究、次世代医療基盤法第25条下にかかるデータ融通の検討を実施した。

A. 研究目的

健康医療ビッグデータをこれまでの統計解析手法に加えてAIを適用させることにより、これまでに省みられることのなかった事象に光をあて、医療の質の向上・均てん化・診療支援と医療分野のイノベーションに貢献することが期待されている。このため、要配慮個人情報収集し、匿名加工された医療情報を円滑に利活用する社会的仕組みとして、医療分野の研究開発に資するための次世代医療基盤法が施行された。この次世代医療基盤法の認定事業者から匿名加工医療情報を提供頂いてICT・AI技術を利用した研究や開発が進展することが期待されている。

しかし、認定事業者は稼働したばかりであり、医学研究に用いられている事例はまだない。また、機械学習手法は多数の変数を要求する傾向がありながら、匿名加工による変数削減とリスクのトレードオフの関係下に変数の数に制約を課せられる可能性があることから、匿名加工医療情報を用いた機械学習の研究について懐疑的な見方もでてい

本研究ではAI技術を用いた研究に関する技術的課題を明らかにし、認定事業者を利用した研究を加速する施策の提言につなげることを目的とする。

B. 研究方法

本研究班は、2年目の活動として、以下に述べる4分野の取組を行った。

(1) 認定事業者を活用するAI研究・制度環境に関するアンケートの実施

認定事業者の潜在的な利用者である地方自治体、学術団体を選定し、Webアンケートシステムにて実施した。

(2) 海外の医療情報の収集・分析にかかる制度・事業の調査

9月に渡航し、英国で医療情報の収集、リンケージ、研究者のデータ利活用に供している組織として、NH S Digital (England); Scottish Safe Havens; Public Health Scotland; University of the West England; Imperial College London; Small Area Health Statistics Unit; Clinical Practice Research Datalinkを訪問し、聞き取り調査を実施した。ま

たその聞き取り調査及び提供された資料をベースに文献調査を実施した。

(3) 前年度から継続しているAI研究

前年度に構築した、LDI内に設置した機械学習環境を用いてLDIの匿名加工医療情報を分析した。J-MIMOから提供された匿名加工医療情報については東京大学医科学研究所のデータセンター内において解析を実行した。

(4) 次世代医療基盤法第25条下にかかるデータ融通の検討

定例会議、エンジニアWG、契約WG、審査WGを延べ30回開催し、次世代医療基盤法第25条下におけるデータ交換にかかる技術的課題、複数の認定事業者間にまたがる契約、審査の課題について検討し、最終提案にむけて資料を取りまとめた。

(倫理面への配慮)

本研究は、認定事業者より提供された匿名加工医療情報を利活用した研究や、個人情報を収集しないアンケートを実施しており、倫理委員会の付議対象ではない。LDIについては安全なデータセンター内で分析し、匿名加工医療情報を機械学習しその結果のみをLDIの確認のもとセキュリティ便で研究者に送付、確認する体制を構築し、情報セキュリティを確保している。J-MIMO提供の匿名加工医療情報の解析については東京大学において高度に管理されたスパコンおよび、研究者のみがアクセスできる、施設された部屋に設置されたPCのみを用いて行った。

C. 研究結果

(1) 認定事業者を活用するAI研究・制度環境に関するアンケートの結果

アンケート依頼を送付した地方公共団体の部署は539部署となった。それらの部署からの総回答数461件中有効回答233件(有効回答率51.5%)であり、アンケート回答率は43.2%であった。31学術団体の会員からは総回答数713件があり、そのうち有効回答として474件(有効回答率66.5%)が得られた。全体的に

次世代医療基盤法・認定事業者に対する認知度は低

て、認定事業者の認知度の向上と認定事業者の利活用を推進するという2つの施策を推進する必要性が見いだされた。

(2) 海外の医療情報の収集・分析にかかる制度・事業
英国において、匿名加工情報を研究者に提供する代わりに、データ加工、分析環境も内包した**Trusted Research Environment(TRE)**というコンセプトを推進しており、TREによってはリモートアクセスも許可していることが確認された。このTREの運用にあたって英国のOffice for National Statisticsが考案した**Five Safe Models**、すなわち”安全な利用者”、”安全なプロジェクト”、”安全な施設/環境”、”安全なデータ”、”安全な分析結果”という5つの安全管理対策の軸にそった事業展開は、わが国でも認定事業者にTREのコンセプトを導入し事業拡大する際の参考になると思われる。次に、データのリンケージにあたって、直接識別子による名寄せとデータそのもののリンクの役務を**Indexer**と**Linker**という独立した組織に担わせることで個人情報露出を最低限におさえた秘匿性の高い運用を実現していることが確認された。本邦が学ぶべき点として、**TRE**、**Five Safe Models**、**Indexer**と**Linker**の分離、利用者への教育体制の拡充が確認された。

(3) AI研究

a) シーケンス解析

オーダーリング情報を対象にシーケンシャルパターンマイニングを実施、頻出医療指示パターン(SPM)の抽出を行う。そして、これらの最長共通部分を導出すると典型的な医療指示パターンを導出できている。複数の医療機関の頻出医療指示パターンを導出し比較することで、医療機関を越えて共通する指示パターンとバリエーションを抽出できることが確認された。

b) 敗血症の予測モデル

入院後の敗血症モデルを予測するための機械学習モデルを開発し、これまでに開発したモデルよりも高い指標(AUC : 0.769、F 値 : 0.528、適合率 : 0.536、再現率 : 0.523) を実現し、多施設から提供された大規模なデータを用いることのメリットが確認された。

c) (ϵ ,k)-乱択匿名化の開発

k-匿名化と差分プライバシー双方の要求を満たす新しい匿名化手法を開発し、EMアルゴリズムと組み合わせることにより高いプライバシーレベルを保ったまま、統計量を高精度に維持できる可能性を提示した。

(4)次世代医療基盤法第25条下にかかるデータ融通の検討

審査WGでのとりまとめとして、複数の認定事業者にわたる申請は最初に申請を受理した認定事業者でワンストップ対応とし、その事業者が審査とリスク評価の審査結果を共有していく方向性を確認した。事業者間の申請書類、共有する項目について突合処理、検討を継続中である。契約WGでは、認定事業者間の医療情報の提供に関する契約書ひな型案を完成させた。エンジニアWGでは、データ授受の形式、手段、媒体、データの統合時のクレンジングの責任分界点、文字コードの扱いについて合意を形成した。

かった。主成分分析、順序カテゴリカル解析を通し

(5)結論

諸般の事情により、LDIの匿名加工医療情報を用いた研究が中断し、LDIとの新規の契約も停止したため、LDIの匿名加工医療情報を用いた研究と次世代医療基盤法第25条下のデータの名寄せの融通にかける検討に関して研究期間を延長することとなった。結果的には4種類のAIにかかる研究を実施し、1つは完了して国際学会の発表を実現し、当初の想定より研究範囲が広がっている。来年度に持ち越された次世代医療基盤法第25条下のデータ融通の検討、認定事業者の潜在的利用者へのアンケート、海外事例の調査を総合的に評価し、最終的に認定事業者にかかる厚生労働省行政の有意義な提言につなげていきたい。

D. 健康危険情報

本研究はリアルワールドデータ(日々の診療行為から発生したデータや、診療報酬、健診データ)の二次利用にもとづいた研究であり、侵襲性のある活動はない。また、次世代医療基盤法の認定事業者の匿名加工医療情報を用いる研究であるため、人を対象とする生命科学・医学系研究に関する倫理指針の対象外である。

認定事業者の潜在的利用者に対する認定事業者の認識にかかるアンケート調査についても、代表研究者が所属する愛媛大学において指针对象外でありIRBの付議対象とはならないという判断を受けている。

E. 研究発表

1) 論文発表

1.Akito Yamamoto, Eizen Kimura, and Tetsuo Shibuya, (ϵ , k)-Randomized Anonymization: ϵ -Differentially Private Data Sharing with k-Anonymity, In Proceedings of the 16th International Joint Conference on Bio medical Engineering Systems and Technologies - HEALTHINF, pp. 287-297.

1.木村映善、高田春樹、荒木賢二、黒田知宏、星佳芳、佐々木香織、伊藤伸介、次世代医療基盤法の認定事業者の認知状況、活用への意識に関する社会調査(投稿準備中)

2) 学会発表

1.Tetsuo Shibuya, "Privacy preserving technologies toward safe medical science", The 47th Annual Meeting of the Japanese Society for Investigative Dermatology (JSID), December 2-4, 2022.

1.木村映善、佐々木香織、伊藤伸介「英国における行政記録情報と医療情報をリンケージした利活用状況について～英国における機械学習での利用を前提とした医療情報の利活用状況の調査から～」『大規模データの公開におけるプライバシー保護の理論と応用』公的統計マイクロデータ研究コンソーシアム/統計数理研究所 共済 2022年12月9日

F. 知的財産権の出願・登録状況 該当なし

令和4年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)
 分担研究年度終了報告書

大規模データの利活用研究の加速のための研究

研究分担者 渋谷 哲朗 (東京大学医科学研究所 教授)

研究要旨

本研究では AI 技術を用いた研究に関する技術的課題を明らかにし、認定事業者を利用した研究を加速する施策の提言につなげることをめざし、実際に事業者より提供された匿名加工医療情報を用いた、差分プライバシーの観点から匿名加工医療情報の AI 研究の可能性の評価を行った。

A. 研究目的

健康医療ビッグデータをこれまでの統計解析手法に加えてAIを適用させることにより、これまでに省みられることのなかった事象に光をあて、医療の質の向上・均てん化・診療支援と医療分野のイノベーションに貢献することが期待されている。このため、要配慮個人情報収集し、匿名加工された医療情報を円滑に利活用する社会的仕組みとして、医療分野の研究開発に資するための次世代医療基盤法が施行された。この次世代医療基盤法の認定事業者から匿名加工医療情報を提供頂いてICT・AI技術を利用した研究や開発が進展することが期待されている。

しかし、認定事業者は稼働したばかりであり、医学研究に用いられている事例はまだない。また、機械学習手法は多数の変数を要求する傾向がありながら、匿名加工による変数削減とリスクのトレードオフの関係下に変数の数に制約を課せられる可能性があることから、匿名加工医療情報を用いた機械学習の研究について懐疑的な見方もでてい

本研究ではAI技術を用いた研究に関する技術的課題を明らかにし、認定事業者を利用した研究を加速する施策の提言につなげることを目的とする。

B. 研究方法

第二号事業者より提供された匿名加工医療情報を用いて、差分プライバシーの観点から匿名加工医療情報のAI研究の可能性の評価を行うことで、提供された情報の評価を行った。

(倫理面への配慮)

本研究は、研究に用いられた情報は匿名加工情報であり、個人情報を明らかにすることのない研究である。なお、提供されたデータの解析には、高度に管理されたスパコンおよび、研究者のみがアクセスできる、施錠された部屋に設置されたPCのみを用いて行った。

C. 研究結果

健康医療データの利用におけるデータ保護の主要な概念としてk-匿名化と差分プライバシーがある。今回提供されたデータはすでにk-匿名化されたデータではあるが、差分プライバシーの観点では厳密には保護されたデータであるわけではなく、差分プライ

バシー保護を要求する国などで用いることはできない。このため、我々は、この2つの概念の双方を満たすデータ加工の方法として、(e,k)-乱択匿名化とよぶ新しい手法を開発した。この手法はk-匿名化と差分プライバシーの双方の要求を満たす新しい手法であり、EMアルゴリズムと組み合わせることにより、高いプライバシーレベルを保ったまま、高精度な人工知能解析も行えることを実験を通して示した[1]。

この研究によって、差分プライバシー加工をk-匿名化に加えて行う新しい匿名化手法を開発することに成功し、将来のデータの利活用に向けた新しい視点を見出すことができた。また、この研究は提供データを用いて検証を行ったものであり、今回の提供データが実際の差分プライバシー研究および人工知能研究に活用可能な非常に有用性の高いものであることも明らかになった。

D. 研究発表

1. Akito Yamamoto, Eizen Kimura, and Tetsuo Shibuya, (e, k)-Randomized Anonymization: ϵ -Differentially Private Data Sharing with k-Anonymity, In Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies - HEALTHINF, pp. 287-297. (本論文に関連して、筆頭著者の山本章人は令和4年度東京大学総長賞を受賞)

E. 学会発表

1. Tetsuo Shibuya, "Privacy preserving technologies toward safe medical science", The 47th Annual Meeting of the Japanese Society for Investigative Dermatology (JSID), December 2-4, 2022.

F. 知的財産権の出願・登録状況
 該当なし

令和 4 年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)
分担研究年度終了報告書

千年カルテの匿名加工医療情報を利用した
入院後敗血症の予測モデルの開発における課題と対策の検討

研究分担者 松田 敦義 (株式会社ログビー)

研究要旨

千年カルテの匿名化された医療データを使用して、入院後の敗血症を予測するための機械学習モデルを開発した。本研究グループの 2020 年度の研究と今回の研究で、ランダムフォレストを用いた予測モデルの評価指標を比較したところ、本研究の評価指標 (AUC : 0.769、F 値 : 0.528、適合率 : 0.536、再現率 : 0.523) がいずれも過去の研究の指標よりも高かった。千年カルテの匿名化された医療データの安全性を担保しながら機械学習モデルを開発するための現状のプロセスの問題点と、その解決策について考察した。問題点として、機械学習モデルの精度と、研究に要する時間の 2 つの観点をあげた。この問題に対し、プロセスの改善策として、テスト環境の準備とデータ集計または機械学習モデル開発のためのテンプレートの提供を提案する。

A. 研究目的

近年、次世代医療基盤法に基づく匿名加工医療情報の活用が注目されている。医薬品の使用による安全性・有効性の研究開発や、患者への個別化医療の提供などを目的とした取り組みが始まっており、期待が高まっている。しかし、その安全性と、AI 活用などにおける有用性についての知見は少ない。

千年カルテプロジェクト (以下、千年カルテと表記) は、2015 年に AMED 研究公募事業に採択された全国共同利用型国際標準化健康・医療情報の収集及び利活用に関する研究である¹⁾。一般社団法人ライフデータイニシアティブ (以下、LDI と表記) は、次世代医療基盤法の認定匿名加工医療情報作成事業者として国から認可を受けた、千年カルテの EHR (Electronic Health Record) を活用して医療の質・効率性や臨床研究等の研究開発に貢献することを目的とした法人である²⁾。

本研究グループは、これまで機械学習と説明可能な AI を用いて、入院後合併症の発症予測と特微量の寄与度の可視化等の研究を行ってきた³⁾。

本研究では、愛媛大学が実施する「認定匿名加工医療情報作成事業者が保有する医療情報を活用した、匿名加工医療情報の作成に依らない AI 研究の実現可能性の検討」研究の一環として、千年カルテの匿名加工医療情報のデータを活用し、LDI の定める安全性を担保するためのプロセス

に従いながら、機械学習モデルの開発を行い、その課題と対策の検討を行った。

機械学習による予測モデル開発では、学習に用いる特徴量の探索やデータの前処理を行うため、データを集計、確認しながら方針を定めることが多い。一方で、千年カルテのデータを利用するに当たり、安全性担保のため、開発プロセスにおいては以下の条件を守る必要があった。

1, 個票データを確認することはできないため、統計データのみを抽出する。

2, データベースを直接参照することはできないため、LDI を通して集計結果やエラーなどのヒアリングをする必要がある。

本研究では、上記のような条件に従い千年カルテの安全性を担保しながら、機械学習モデルの開発を行い、そのプロセスの課題の抽出と対策の検討を行うことを目的とした。

機械学習モデルとしては、入院時の患者の状態を元に、入院後の敗血症の発症を予測するモデルを開発した。機械学習モデルの精度評価において、本研究グループが 2020 年度の第 40 回医療情報学連合大会で発表した、宮崎大学医学部附属病院のデータを用いた研究結果との比較を行った。

B. 研究方法

1. 開発プロセス

以下に示す 2 つのステップ A, B で、機械学習モデルの開発を進めた。

A, 宮崎大学医学部附属病院のデータセンター（以下、宮崎大学環境と表記）に、千年カルテの環境と近いデータベース、解析サーバーを構築し、そこでコンテナ型の仮想環境 (Docker) を作り、宮崎大学医学部附属病院のデータを用いて前処理、学習と推論を行うプログラムの開発を行った。宮崎大学環境での機械学習モデル開発の概要を図 1 に示す。

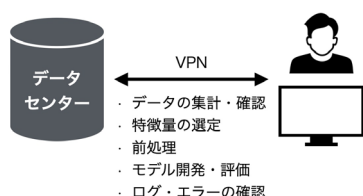


図 1 宮崎大学環境での機械学習モデル開発

B, 千年カルテの本研究用コンピューター（以下、千年カルテ環境と表記）に、A で作成した Docker イメージを持ち込んでコンテナ化し、LDI と連携して試行錯誤しながら開発したプログラムを改修し、機械学習モデルの開発を行った。千年カルテ環境における開発では、以下の手順を繰り返して進めた。

- 1, プログラムと実行手順を LDI に提供
- 2, LDI によるプログラム実行とエラー等のフィードバック
- 3, プログラムを改修し LDI に提供
- 4, LDI 担当者がデータセンターに入室してプログラムを実行し、データ統計やログを取得

千年カルテ環境での機械学習モデル開発の概要を図 2 に示す。

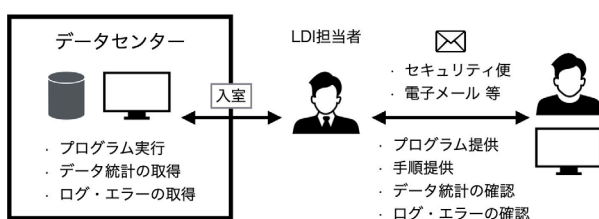


図 2 千年カルテ環境での機械学習モデル開発

最終的に開発した機械学習モデルの評価結果と特徴量重要度を LDI よりセキュリティ便で受領した。

2. 対象データ

千年カルテに登録された 20 の医療機関の 2015～2021 年度の DPC を使用した。また本研究での対象とする患者グループの入院時疾患は ICD10 分類における C00-D48 “新生物<腫瘍>” とし、予測対象の入院後合併症を ICD10 分類における

分類表記で“敗血症”を含む疾病を対象とした。予測に用いる特徴量として、DPC の年齢や体重などの基本情報、主傷病や入院時併存症等の診断情報を使用した。使用した特徴量は全部で 166 個であった。

3. 前処理

機械学習に用いるデータは入院単位のデータとなるように整形した。また予測の正当性を守るために、入院時に敗血症を発症している患者は除いた。その他の前処理として、カテゴリデータのフラグ化、欠損のあるデータの削除等を行った。結果として、合計 71,433 件のデータとなった。

以上の処理の後、データ全体を予測モデルの学習に 75%、評価に 25%となるように分割した。学習データについて、データ数が敗血症の発症：非発症 = 1：20 になるように非発症データをダウンサンプリングした。

4. 予測モデル

本研究では、プログラムの実装がしやすく、予測に寄与した特徴量の重要度を得ることができる Random Forests⁵⁾（以下、RF と表記）を機械学習モデルとして用いた。前処理済のデータで学習させ、各ハイパーパラメータは交差検証を行い決定した。

（倫理面への配慮）

本研究は、千年カルテの利用目的等審査委員会⁴⁾による審査を受け、承認された上で研究を進めた。

C. 研究結果

1. 開発プロセス

千年カルテ環境における開発の手順を進めるにあたり、LDI との電子メールのやり取りは 100 回近くに及んだ。やり取りを通して、時間がかかったプロセスを表 1 に示す。定性的に時間がかかったと考えられる順番で記載した。

各プロセスの中で、「プログラム実行で発生したエラーの確認と対応」に多く時間がかかった。理由としては、エラー時のログについて、LDI にて内容の確認を行い、電子メールにて伝達をする運用であったため、エラーの全体像の把握が難しく、推測してプログラムを直す必要があったことが大きい。また、LDI 担当者もデータセンターのセキュリティルームに入室した上でプログラムを実行し、データ統計やログの取得を行う必要があったため、LDI の運用プロセスにおいても時間がかかった。

表 1 時間がかかった開発プロセス

開発プロセス	
1	プログラム実行で発生したエラーの確認と対応
2	データ授受等、研究の進め方の確認
3	データベース内のテーブルの概要確認と、対象テーブルの修正
4	データベース接続情報の確認と、接続情報の修正
5	Docker イメージでインストールされるライブラリの確認と承認
6	千年カルテから取得するデータ種類の確認と承認

2. 予測モデルの精度評価

予測精度の評価指標として AUC (Area Under the Curve) 及び F 値 (適合率と再現率の調和平均)、適合率、再現率を用いた。

表 2 に、本研究グループが 2020 年度に発表した RF を用いた敗血症の予測モデルと、本研究における敗血症の予測モデルの評価指標の比較を示す。各指標の値は、敗血症を発症するケースとしないケースの平均をとったものである。いずれの評価指標においても、本研究の精度が高い結果となった。

ただし、2020 年度の研究と本研究では用いた特徴量やデータの母数などの前提条件が異なるので、一概に単純比較はできない。2020 年度の研究では、DPC だけでなく DWH からバイタル、検体検査、実施オーダー情報 (薬剤処方、手術、放射線) 等も利用し、特徴量としては 358 個を用いたのに対し、本研究では DPC のみから 166 個の特徴量を用いた。また、2020 年度の研究では宮崎大学医学部附属病院のみの前処理済データ 13,527 件を利用したのに対し、本研究では千年カルテの 20 の医療機関の前処理済データ 71,433 件を用いた。

2020 年度の宮崎大学医学部附属病院のみのデータ数に対して、本研究の 20 の医療機関のデータ数が 5 倍程度と、施設数の比率の 20 倍に対して小さかった 1 つの要因として、2020 年度では前処理で欠損値のデータ補完を行ったのに対し、本研究では欠損があるデータはすべて削除したことが考えられる。

表 2 2020 年度の研究と本研究の評価指標の比較

研究	AUC	F 値	適合率	再現率
----	-----	-----	-----	-----

2020 年度	0.727	0.424	0.374	0.490
本研究	0.769	0.528	0.536	0.523

表 3 に、本研究において、敗血症の発症なし、発症ありで分けた場合の評価指標を示す。各指標において発症なしの値は非常に大きい、発症ありの各指標、特に再現率の値が小さい事がわかる。このことから、表 2 の本研究の評価指標は発症なしの評価結果により値が高められているといえる。発症ありの精度を高めることが今後の課題である。

表 3 本研究における敗血症の発症別の評価指標

敗血症	F 値	適合率	再現率
発症なし	0.997	0.996	0.998
発症あり	0.058	0.075	0.048

3. 予測モデルの特徴量重要度

本研究の予測モデル (RF) の特徴量重要度について、重要度が大きい順で上位 10 個を表 4 に示す。これを見ると体重、身長、年齢が上位に来ていることがわかる。2020 年度の研究で用いたバイタル、検体検査、実施オーダー情報 (薬剤処方、手術、放射線) 等の特徴量を追加すると、精度向上に寄与する重要な特徴量に変化するのではないかと期待される。

表 4 本研究における予測モデルの特徴量重要度

	特徴量	重要度
1	体重	0.148074051
2	身長	0.130524826
3	年齢	0.075204027
4	併存症_真菌症	0.036117323
5	併存症_原虫疾患	0.034309911
6	併存症_悪性新生物<腫瘍>	0.025880966
7	併存症_高血圧性疾患	0.025158698
8	併存症_糖尿病	0.022879724
9	併存症_食道、胃、および十二指腸の疾患	0.022592592
10	性別	0.021793955

4. 開発プロセスの課題

本研究での千年カルテ環境の開発プロセスにおいて、機械学習モデルの精度の観点と、研究にかかる時間の観点で、それぞれ課題が考えられる。機械学習モデルの精度の観点では、個票データ

の確認ができないため、データを確認しながら特微量の内容（例えば検査値の単位やノイズによる影響など）や表記ゆれなどの確認ができないため、データが正しい前提で開発を行う必要があり、実際に何かデータの問題を内包していた場合には、精度が落ちる等の問題が考えられる。

研究にかかる時間の観点では、図2に示したように、データの統計量や結果はLDIを通して確認する必要があるため、データを直接確認しながら機械学習モデルの開発を行うよりも時間が多くかかる問題がある。教師あり機械学習モデルを開発する上では、様々なデータを集計し、適切な特微量を選定するプロセスが必要である。本研究では、宮崎大学環境にて選定した特微量をそのまま千年カルテ環境でも用いたため、表1の「時間がかかった開発プロセス」には含めていないが、より精度の高い機械学習モデルの開発をするためには、特微量選定に多く時間がかかると考えられる。

上記の問題を軽減するためには、千年カルテ環境で使おうとしているデータに精通した研究者が、手戻りやLDIとのやり取りができるだけ少なくなるように、予め網羅的なデータ統計やログ出力の設計を行い、一度のやり取りで多く情報を取得できるように進めることが必要である。

5. 開発プロセスの課題への対策

開発プロセスの課題への対策として、以下の2点が考えられる。

1, 千年カルテ環境と同様の検証環境を用意する。

本研究では、予め宮崎大学環境にて機械学習モデルを開発した上で、千年カルテ環境での開発に臨んだ。このプロセスがなければ、本研究を進めることは難しかったと考えられる。もし千年カルテにて同様の検証環境が用意されれば、まずそこで研究の方針を立て設計することができ、本番環境で手戻りが少なくなると考えられる。

2, データ集計や機械学習モデルを開発するテンプレートを提供する。

千年カルテの利用者は、LDIに集計したい項目や、開発したい機械学習モデルの種類などを提示すれば、結果を得られるという状態が理想的である。網羅的なテンプレートを提供するには、相当の時間がかかると想像されるため、例えば過去の

研究で行われたプロセスなどをテンプレート化していくなど、少しずつ進める形が現実的ではないかと考えられる。

6. 多施設の匿名加工医療情報を用いる効果

2020年度の研究と本研究の評価指標の比較から、多施設のデータを利用することにより、単一の医療機関のデータを用いた機械学習モデルよりも予測精度を向上できる可能性が示唆された。本研究の結果は特微量の選定や欠損値の補完などがきちんとできていない上で得られたもので、学習前のプロセスをより精密に行うことで、更に精度を高めていく事ができると考えられる。

多施設のデータを用いてデータの母数を増やす事により、これまで難しかった大規模な予測モデルの開発や、多施設間での臨床データの比較、希少疾患の分析への活用などの用途が期待される。

D. 健康危険情報

総括研究年度終了報告書にまとめて記入する。

E. 研究発表

該当なし

F. 知的財産権の出願・登録状況

該当なし

参考文献

- 1) 千年カルテプロジェクト概要. 日本医療ネットワーク協会.
[<https://www.gehr.jp/about/index.html> (cited 2023-Apr-23)]
- 2) 認定匿名加工医療情報作成事業者について. 次世代医療基盤法. 一般社団法人ライフデータイニシアティブ
[<https://www.ldi.or.jp/law> (cited 2023-Apr-23)]
- 3) 松田敦義. et al. 複数の入院後合併症に対する時系列予測モデルの開発と説明可能なAIを用いたリスク要因の比較. 医療情報学会, 2020.
[https://jglobal.jst.go.jp/detail?JGLOBAL_ID=202002212479945038 (cited 2023-Apr-23)]
- 4) 利用目的等審査委員会の公表. 一般社団法人ライフデータイニシアティブ
[<https://www.ldi.or.jp/committee> (cited 2023-Apr-23)]
- 5) Breiman L. Random Forests. Machine Learning 2001 ; 45(1) : 5–32.

令和4年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）
分担研究年度終了報告書
大規模データの利活用研究の加速のための研究

研究分担者 佐々木香織（北海道公立大学法人札幌医科大学 医療人育成センター 教授）

研究要旨

2018年より次世代医療基盤法が施行され、新たに設立された認定事業者を通じた医療情報の利活用の促進が期待されている。しかし機械学習の利用や行政情報との連携は具体的には未整備なため、これらを視野に入れた認定事業者のあり方について考慮する必要があるだろう。そこで本研究は個人情報保護と市民社会の関係性において類似点が日本と多く、且つ機械学習と行政情報の利活用を含めた医療情報利活用の制度整備が進んでいる英国の事例から学ぶ。そこから認定事業者にかかる施策提言に発展させる予定である。

A. 研究目的

2018年より次世代医療基盤法が施行され、医療情報の利活用の促進が期待されている。同法で想定している利活用とは、認定事業者が提供する匿名加工医療情報を用いた研究や開発を進展することである。しかし匿名加工情報を用いた機械学習においては、懐疑的な見方もでている。機械学習手法は多数の変数を要求する傾向があるが、匿名加工による変数削減とリスクのトレードオフの関係下で変数の数に制約を課せられる可能性があり、その様な要求に応えることが難しいと推測されるからである。更に認定事業者は行政情報も扱える設計だが、医療情報との連携は未整備である。したがって機械学習と行政情報の利活用を視野に入れた発展が必要と考えられる。

海外でも医療情報の機械学習や行政情報と連携した利活用への期待は高いが、各国で市民社会における個人情報への態度が異なるため、様々な体制が整備されている。英国は個人情報保護と市民社会の関係性が日本と類似する点が多い一方、機械学習を含めた医療情報利活用の体制整備が進んでいる。そこで本分担研究は、英国の事例から本邦が学ぶべき点を探究することを目的とする。

B. 研究方法

本研究は文献調査と令和4年9月に英国の関連団体 NHS Digital (England); Scottish Safe Havens; Public Health Scotland; University of the West England; Imperial College London; Small Area Health Statistics Unit; Clinical Practice Research Datalink に実施した聞き取り調査からなる。

（倫理面への配慮）

個人情報・医療情報を扱わないため、該当なし

C. 研究結果

英国はEngland, Wales, Scotland, Northern

Ireland の連合王国であるため、公的医療サービスは四王国に分かれて運営される。そのうち本研究は紙幅の関係上Scotlandの医療情報利活用を支える公的機関(i.e., Scottish Safe Haven)に焦点をあてる。なぜならScotlandが全英で利活用の体制整備を最も発展させたからである。また一次医療圏のデータ利活用を扱うCPRDにも触れる。全英レベルの組織で自主独立運営という日本の認定事業者との類似性と、近年設立の他組織と異なり30年の歴史という経験知から、参考にすべき点が多いからである。

(1) Scottish Safe Haven*

英国では主に各王国のNational Health Service(NHS)という公的機関を通じて、医療サービスが提供される。Scotland王国のNHS Scotlandでは、医療情報の利活用の体制を整備するにあたり、2008年に組織された官民連携の研究グループScottish Informatics Programme(SHIP)が制作したBlueprintという報告書**にもとづいて、新たにSafe Havenという組織を設立させた。それはNHSの編成に倣い国と地方に分かれている。具体的には、国家レベル組織のNHS Scotland内にあるInformatic Service Division下にNational Safe Havenを配置し、Aberdeen, Dundee, Edinburgh, Glasgowの四地域にある大学附属病院を擁する地方単位のNHSには、四つの地域Safe Havenを配置している。

ここで注目すべき点は二つある。一つがTrusted Research Environment (TRE)という環境の整備だ。

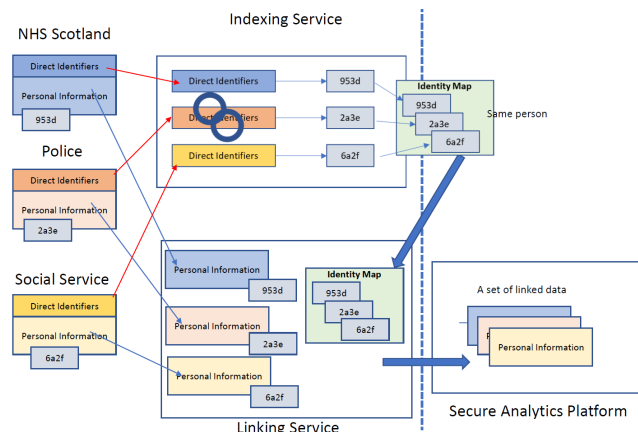
それはScottish Safe Havenが匿名加工情報を研究者に提供するのではなく、TREという研究目的の安全なIT環境と仮名化した医療情報の二つを提供し、TREという環境下で提供したデータを用いた研究を行ってもらうサービス体制である。それにより機械学習を含めて匿名加工情報では研究の深化や発展が難しい案件を扱うことが可能となる。

しかし仮名情報の使用を鑑み、TREの安全性を盤石にする必要がある。そこでScottish Safe Havenでは英国のOffice for National Statisticsが考案したFive Safe Modelsを発展的に採用し、安全性を担保している。その概要*は以下の通りだ。

1. 安全な利用者(Safe People): データ利用は定められたガイドラインと規則に従うことができる研究者に限る。
2. 安全なプロジェクト(Safe projects): プロジェクトは公共の利益に資するもので、データを適切に扱うことを確約するものに限る。また患者と市民参画、意思決定過程とデータの利用に関する透明性を担保できるものに限る。
3. 安全な施設/環境(Safe settings): データが安全に保管され、利用者が安全にデータ分析・解析できるTREを提供。且つ利用者の作業を追跡して点検ができるように設定。
4. 安全なデータ(Safe data): 研究者が分析する過程でデータ対象者を直接特定できるような情報を削除するなどの保護。それによりEU離脱前に批准した通称GDPR (General Data Protection Regulation = 一般データ保護規則) に則った形でデータ秘匿性を担保。
5. 安全な分析結果(Safe outputs): データや成果物の持出しは、個人情報漏洩や特定するリスク点検後に許可されたものだけに限定。

第二は、上述のBlueprintに基づいたデータ突合にかかる制度設計である。各Safe Havenでは突合にあたりIndexerとLinkerの組織を分けることで、個人情報保護を担保している。しかし行政機関など所管が異なるデータの突合の際にはSafe Haven外の第三者機関とも連携 (e.g., National Record of Scotland での Indexing; Public Health England 内のeDRISでのlinking) している。下図***のように、識別情報を扱う機関(Indexer)と個人の状況に関する情報を扱う機関(Linker)が別個なため、秘匿性を極めて高くした設計となっている。

* 本セクションはScottish Safe HavenのコアメンバーであるEmily Jefferson 教授への聞き取り(2022年9月15日)と、その際に提示された彼女の2021年度University of Dundee 大学院講義のス



ライドをもとにしている

** SHIP (2012) Blueprint for healthcare research in Scotland. [Online] http://www.scotship.ac.uk/sites/default/files/Reports/SHIP_BLUEPRINT_DOCUMENT_final_100712.pdf

*** 上註*の講義スライドを、研究代表者の木村が構成し筆者が一部改変

(2) Clinical Practice Research Datalink (CPRD)****

GPRDもScottish Safe Havenと同様の背景でTREへの移行が進み、ほぼ同様のFive Safe Modelを採用している。それに伴い適切なTREの構築と利用者への教育を検討課題として挙げている。また検討の際に特に問題視されているのが、NPOとして受益者負担で運営されている一方で、データの収集・保守・管理や匿名加工といった作業には無料か少額の費用で済むと考えている人が一般市民のみならずCPRDの利用者の中にも一定数いるため、費用への不満がCPRDの設立から三十数年を経ても散見される点である。そこで以下のバランスが今後の発展に重要だとされている。1. 公益に資する研究を行うニーズに、どこまで応えて研究環境を提供するか、2. データ提供の対価として、どの様な価格設定が適切で研究者が負担可能な価格ほどの位か。

****本セクションは、CPRD Deputy Director Dr Puja Mylesへの聞き取り(2022年9月23日)をもとにまとめてある。

考察

本邦が学ぶべき点は以下の四点に収斂するだろう。

- TRE整備のあり方 (ニーズと価格の考慮を含む)
- Five Safe Modelsのあり方
- IndexerとLinkerの分離の可否とそのあり方
- 利用者への教育体制の検討

D. 健康危険情報 該当なし

E. 研究発表（学会発表）

1.木村映善, 佐々木香織, 伊藤伸介「英国における行政記録情報と医療情報をリンケージした利活用状況について～英国における機械学習での利用を前提とした医療情報の利活用状況の調査から～」
『大規模データの公開におけるプライバシー保護の理論と応用』公的統計マイクロデータ研究コンソーシアム/統計数理研究所 共済 2022年12月9日

F. 知的財産権の出願・登録状況

該当なし

令和 4 年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)
 分担研究年度終了報告書

千年カルテの匿名加工医療情報を利用した
 他医療機関の電子カルテに対するシーケンス解析に関する研究

研究分担者 横田 治夫 (東京工業大学)

研究要旨

複数医療機関における頻出医療指示パターンを比較し、医療機関間の医療指示パターンの共通点や相違点を明らかにすることで、自医療機関の特徴を確認し、他医療機関の治療パターンを参照した医療行為の改善等に繋げることが期待される。しかし、比較対象の医療機関数が多くなると共通部分が減り、医療機関の特徴を正確に把握することが難しくなる。そこで、パターンの共通部分を基にした頻出医療指示パターン間の距離を定義し、クラスタリングによって対象となる医療機関群をいくつかクラスタに分類し、そのクラスタ内で比較を行うことで特徴を把握しやすくする手法を提案する。本年度はオプトアウトされている 20 医療機関のデータに適用し、17 医療機関から頻出医療支持パターンを抽出できた。来年度はこれらのパターンに対し医療機関間の頻出医療支持パターンの共通点・差異点を示し、評価を行う。

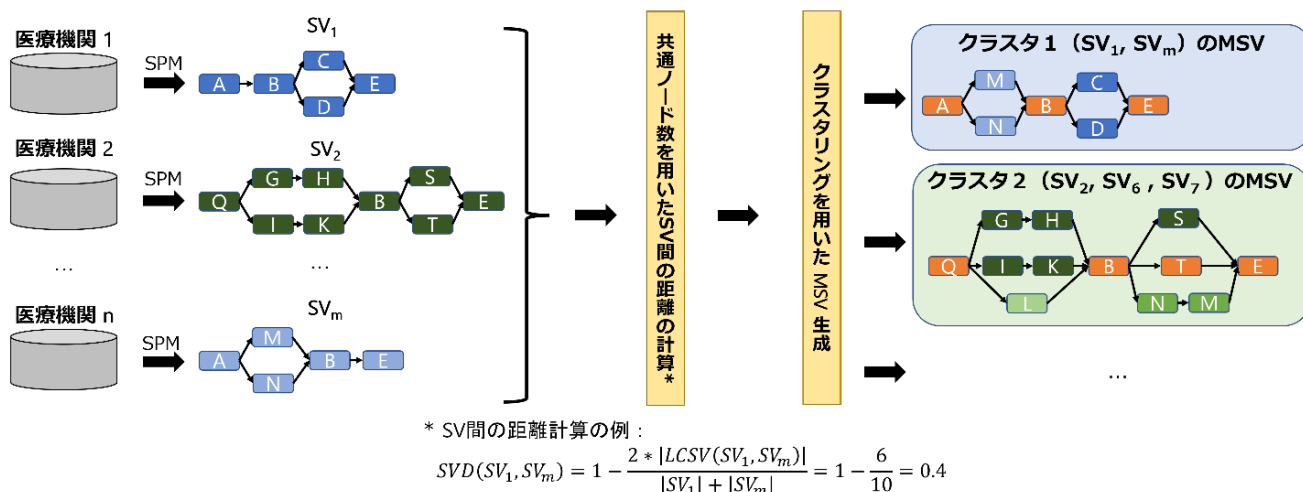


図 1：複数医療機関の電子カルテに対する頻出医療支持パターンの解析

A. 研究目的

複数医療機関における頻出医療指示パターンを比較し、医療機関間の医療指示パターンの共通点や相違点を明らかにすることで、自医療機関の特徴を確認し、他医療機関の治療パターンを参照した医療行為の改善等に繋げることが期待される。しかし、比較対象の医療機関数が多くなると共通部分が減り、医療機関の特徴を正確に把握することが難しくなる。そこで、パターンの共通部分を基にした頻出医療指示パターン間の距離を定義し、クラスタリングによって対象となる医療機関群をいくつかクラスタに分類

し、そのクラスタ内で比較を行うことで特徴を把握しやすくする。

B. 研究方法

1. 背景

これまで電子カルテ中の医療指示のシーケンス解析を行い、検体検査項目をクラスタリングして検査タイプを抽出し、検査タイプと検査結果のシーケンスから次の検査項目を推薦する手法を開発してきた 2)。また、医療オプションとなるバリエーションの安全性、効率性の評価や分岐理由の推定の提案と評価を行った 3)。さらに、

2 医療機関の医療指示のパターンの共通点・差異点を見せる手法も提案した4)。これらの研究の知見を活かし、ライフデータイニシアティブ (LDI) が収集した複数医療機関における医療指示を解析する手法の研究を行う。

2. 開発プロセス

千年カルテの本研究用コンピューター（以下、千年カルテ環境と表記）に、Docker イメージを持ち込んでコンテナ化し、LDI と連携して試行錯誤しながら開発したプログラムを改修/・開発を行った。千年カルテ環境における開発では、以下の手順を繰り返して進めた。

- 1, プログラムと実行手順を LDI に提供
- 2, LDI によるプログラム実行とエラー等のフィードバック
- 3, プログラムを改修し LDI に提供
- 4, LDI 担当者がデータセンターに入室してプログラムを実行し、データ統計やログを取得

最終的に開発したプログラムの評価結果を LDI よりセキュリティ便で受領した。

3. 対象データ

千年カルテに登録された 20 の医療機関の 2015～2021 年度のデータを使用した。また本研究では入院中に経皮冠動脈インターベンション (PCI) を受けた患者を対象とした。

4. 提案手法

図 1 に示すように、まず各医療機関の医療指示データを対象に、シーケンシャルパターンマイニング (SPM) を行い、頻出医療指示パターンの抽出を行う。ここで頻出医療指示パターンをシーケンスバリエーション (SV) として表し、SV 間の共通部分である最長共通サブシーケンスバリエーション (Longest Common Subsequence Variant, LCSV) を導出する。そして、医療機関間の頻出医療指示パターンの差異を表す距離を LCSV のノード数を用いて定義する。具体的に、2 つの SV の間には LCSV のノード数が多いほど、その SV 間の距離が短くなるように計算する (図 1 の計算例参照)。これを用いて階層的クラスタリングを行うことで医療機関群をいくつかのグループに分類する。さらに、グループ内で LCSV を算出し、併合シーケンスバリエーション (Merged Sequence Variant, MSV) を作成することで、医療機関の頻出医療指示パターンの共通点 (桃色) と差異点を可視化する。

(倫理面への配慮)

本研究は、千年カルテの利用目的等審査委員会 1) による審査を受け、承認された上で研究を進めた。

C. 研究結果

本年度はオプトアウトされている 20 医療機関のデータに SPM を適用し、17 医療機関からそれぞれの頻出医療指示パターンを抽出することができた。来年度はこれらの頻出医療指示パターンに対して、SV の距離に基づいたクラスタリングを行い、医療機関間の頻出医療指示パターンの共通点と差異点を示す MSV を可視化し、評価を行う。

D. 健康危険情報

総括研究年度終了報告書にまとめて記入する。

E. 研究発表

本年度は、学会発表を行わなかった。

F. 知的財産権の出願・登録状況

本年度は、知的財産権の出願・登録を行わなかった。

参考文献

- 1) 利用目的等審査委員会の公表. 一般社団法人ライフデータイニシアティブ

[<https://www.ldi.or.jp/committee> (cited 2023-Apr-23)]

- 2) Le et al., Sequential Pattern Mining for Large Combinable Items with Values for a Set-of-items Recommendation, Proc. 34 IEEE International Symposium on Computer-based Medical Systems (CBMS2021), pp. 56-61, 2021.

- 3) Le et al., Methods for Analyzing Medical-order Sequence Variants in Sequential Pattern Mining for Electronic Medical Record, ACM Transaction on Computing for Healthcare, vol. 4, iss. 1, article No. ;3, pp. 1-28, 2023.

- 4) Li et al., Comparison of Sequence Variants and the Application in Electronic Medical Records, Proc. 33rd International Conference on Database and Expert Systems Applications (DEXA 2022), pp. 117-130, vol. 13427, 2022.

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書 籍 名	出版社名	出版地	出版年	ページ
該当なし							

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
Akito Yamamoto, Eizen Kimura, Tetsuo Shibuya	ϵ -Differentially Private Data Sharing with k -Anonymity, In Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies	HEALTHINF		287-297	2022
木村映善、高田春樹、荒木賢二、黒田知宏、星佳芳、佐々木香織、伊藤伸介	次世代医療基盤法の認定事業者の認知状況、活用への意識に関する社会調査	保健医療科学		投稿準備中	2023

令和 5年 4月 18日

厚生労働大臣 殿

機関名 国立大学法人愛媛大学

所属研究機関長 職 名 大学院医学系研究科長

氏 名 羽藤 直人

次の職員の令和4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。

- 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）
- 研究課題名 大規模データの利活用研究の加速のための研究
- 研究者名 （所属部署・職名）大学院医学系研究科・教授
（氏名・フリガナ）木村 映善・キムラ エイゼン

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）
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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5年4月1日

厚生労働大臣 殿

機関名 宮崎大学

所属研究機関長 職名 医学部長

氏名 菱川 善隆

次の職員の令和4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 大規模データの利活用研究の加速のための研究

3. 研究者名 (所属部署・職名) 宮崎大学医学部附属病院 患者支援センター ・ 研究員

(氏名・フリガナ) 荒木賢二 アラキケンジ

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)
人を対象とする生命科学・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 checked="" 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 checked="" 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 checked="" 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 checked="" 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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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"/>
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"/>

(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 京都大学

所属研究機関長 職名 医学研究科長

氏名 伊佐 正

次の職員の令和4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。

- 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）
- 研究課題名 大規模データの利活用研究の加速のための研究
- 研究者名 （所属部署・職名） 医学研究科・教授
（氏名・フリガナ） 黒田 知宏・クロダ トモヒロ

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）
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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 5年 3月31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 国立保健医療科学院

所属研究機関長 職 名 院長

氏 名 曾根 智史

次の職員の令和4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 大規模データの利活用研究の加速のための研究
3. 研究者名 (所属部署・職名) 研究情報支援研究センター長
(氏名・フリガナ) 星 佳芳 (ホシ ケイカ)

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)
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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 5年 3月 31日

厚生労働大臣 殿

機関名 国立保健医療科学院

所属研究機関長 職名 院長

氏名 曾根 智史

次の職員の令和4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。

- 研究事業名 政策科学総合研究事業（統計情報総合研究事業）
- 研究課題名 大規模データの利活用研究の加速のための研究
- 研究者名（所属部署・職名） 研究情報支援研究センター 特任研究員
（氏名・フリガナ） 水島 洋 ミズシマヒロシ

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）
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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5年4月4日

厚生労働大臣 殿

機関名 国立大学法人東京大学

所属研究機関長 職名 学長

氏名 藤井 輝夫

次の職員の令和4年度 厚生労働科学研究費補助金 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 大規模データの利活用研究の加速のための研究
3. 研究者名 （所属部署・職名） 医科学研究所 ・ 教授
（氏名・フリガナ） 渋谷 哲朗 ・ シブヤ テツオ

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）
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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 5年 4月 11日

厚生労働大臣 殿

機関名 札幌医科大学

所属研究機関長 職名 学長

氏名 山下 敏彦

次の職員の令和4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。

- 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）
- 研究課題名 大規模データの利活用研究の加速のための研究
- 研究者名 （所属部署・職名） 医療人育成センター・教授
（氏名・フリガナ） 佐々木香織・ササキカオリ

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）
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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 type="checkbox"/> 無 <input checked="" type="checkbox"/> （無の場合はその理由：規程の申告要件に該当しないため）
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5年4月13日

厚生労働大臣 殿

機関名 中央大学

所属研究機関長 職名 学長

氏名 河合 久

次の職員の令和4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 大規模データの利活用研究の加速のための研究

3. 研究者名（所属部署・職名） 中央大学経済学部・教授

（氏名・フリガナ） 伊藤 伸介・イトウ シンスケ

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）
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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5年4月21日

厚生労働大臣 殿

機関名 一般財団法人 日本医師会医療情報管理機構

所属研究機関長 職名 代表理事

氏名 茂松 茂人

次の職員の令和4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。

- 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）
- 研究課題名 大規模データの利活用研究の加速のための研究
- 研究者名 （所属部署・職名） 理事
（氏名・フリガナ） 長島 公之・ナガシマ キミュキ

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）
人を対象とする生命科学・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"/>
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"/>
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」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系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。