

## 医用 AI 開発とデータ駆動型医療実現を目的とした、高精度医療リアルワールドデータ生成および統合解析共通プラットフォームの開発研究計画

### ① 計画の概要

基礎医学（社会医学含む）と臨床医学の知を統合し、双方向へ還元する医学医療情報の大規模統合解析と医用人工知能（AI）応用基盤を研究開発する。具体的には A）臨床課題指向型構造化次世代電子カルテシステムの研究開発と実装、各種生体センサーと IoT により生活や介護データを収集管理する個人健康情報管理システム（PHR）の標準化と普及による高品質の医療データ生成システムの構築、B）次世代医療基盤法を活用した個人の健康医療情報を統合する匿名加工化健康医療情報データベース（FHDB）の整備、C）FHDB を活用する医学医療統合データ解析活用センターの構築と医用 AI 応用システムの開発、の 3 本柱からなる。

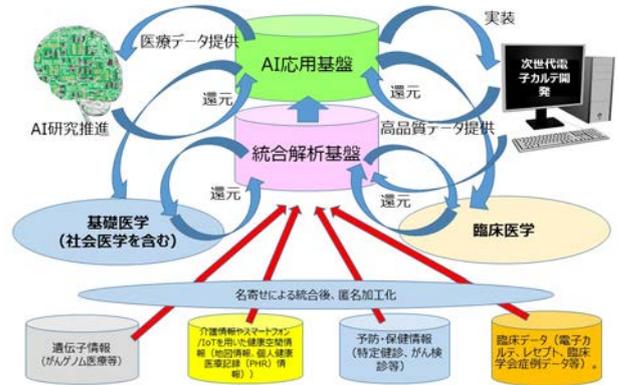


図 1 研究概要

まず臨床データ（電子カルテ、レセプト等）、予防保健情報（特定健診、がん検診等）、PHR 情報（介護情報含む）、遺伝子情報（がんゲノム医療等）を多機関から非匿名データで認定匿名加工事業者が収集し、名寄せ統合後に匿名加工した FHDB を整備する。大規模な匿名加工情報の解析成果を、創薬を含む臨床研究や基礎研究へ還元し、また深層学習等の機械学習を適用した医用 AI 技術開発などの広範な医学医療・医用工学研究を支援しうる解析センターを構築し、安全にアクセスできるサテライトを全国 3 箇所に展開する。同時に、高品質で臨床課題解決指向に構造化された医療データを大規模に生成できる次世代電子カルテを開発実装し、FHDB にデータ提供することにより解析精度やパワーを向上する。

本研究開発により、臨床研究や基礎研究に今まで見えてこなかった課題、あるいは解決が困難であった課題解決の基盤を得る。次世代電子カルテは、エビデンスレベルの高い知を生成しうる基盤のコアであり、医療スタイルにも大きな転換を与える。さらに AI 技術の研究者に医学医療の大規模データを利用できる環境を提供すること自体、医用 AI 研究を飛躍的に推進する。

### ② 学術的な意義

これまで克服できなかった、基礎・臨床医学の領域間、および情報理工学研究領域との間の高い垣根を取り払うことができるのが次世代医療基盤法上で生成するビッグデータである。事業者毎に散在する膨大な保健・医療・介護データを、次世代医療基盤法上で個人の一貫したデータとして統合し、匿名加工化して得た大量のデータを活用し、応用統計技術や機械学習を駆使し、旧来の研究では実現しえない規模と速度で人類に役立つ新たな医学知を量産する基盤を構築することが、本研究の学術的意義の真髄である。そのためには、高品質で構造化された大規模医療データを生成できる次世代型の高精度医療データ生成電子カルテシステムの構築が必要不可欠であり、それは医療スタイルにも大きな転換を与える。膨大な医療 RWD は新たな知を生成し、新たな知は次なる基礎研究の萌芽を促し、さらに多くのデータを生み出す循環をもたらす。これらを通じ、各疾患の悉皆データを小さいタイムラグでバイアスなく収集解析し、極めて信頼性の高いエビデンスを極めて低コスト、短期間で得られる。長い期間と莫大なコストをかけていた臨床アウトカムの評価や薬剤副作用の発生状況の把握、予期しない臨床イベント発生率の解析などが、リアルタイムかつ低コストで実現できる。AI 技術の研究者に大規模医療データ利用が可能な環境を提供すること自体が、医用 AI 研究を飛躍的に推進し、多くのデータ・アナリストと医用 AI 研究者を育成しうる。

### ③ 国内外の動向と当該研究計画の位置づけ

医療 RWD を複数機関で統合活用する試みは、制限された規模やデータ項目種では国内外で進展の緒についているところである。医療情報の国際標準化は日本医療情報学会等が推進している。標準化された手法による臨床症例データベースの構築が始まり、MID-NET 事業は複数施設の医療 RWD を標準化して解析するサービスを実用化した。全国レセプトデータは厚労省が公開を始め、東大と京大にデータ利用センターが作られた。個人データの匿名加工による二次利用は、2018 年に次世代医療基盤法が施行され匿名加工認定事業者の申請が始まった。本計画は、臨床課題指向型構造化高品質データを生成できる次世代電子カルテシステムを開発するとともに、情報の収集・標準化・統合・解析システム開発を行い、臨床課題指向型標準情報基盤を実現し、その利活用プラットフォームを構築するものであるが、診療の場で入力される電子カルテデータだけを対象とするのではなく、生活や介護の場の生体データも含み、高品質で構造化された医療 RWD を先端的融合技術により生成し、そのデータを日常診療だけでなく研究にも提供する点で既存の研究動向を包括し、さらに拡張するという特徴を持つ。

### ④ 実施機関と実施体制

◎：実施組織における総括代表、○：主な実施機関

主な実施機関：◎一般財団法人日本医療情報学会、○九州大学（病院メディカル・インフォメーションセンター）、東京大学大学院医学系研究科（社会医学専攻の関連講座、医学部附属病院）、大阪大学（情報統合医学講座）、○京都大学（病院医療情報企画部）、名古屋大学（病院医療情報部）、帝京大学（医療情報システム研究センター）、東北大学（医療情報学分野）

主な協力機関：一般財団法人医療情報システム開発センター（MEDIS-DC\*）、一般社団法人保健医療情報システム工業会（JAHIS\*）、日本糖尿病学会\*、日本腎臓学会\*、日本クリニカルパス学会\*、一般社団法人診断群分類研究支援機構※、日本臨床疫学会※（\*の機関とは合同委員会等での活動を普段より実施中、※の機関は交渉中）

実施組織における主な機関の役割は、以下を予定している。

日本医療情報学会：全体の統括、高精度医療データ生成システムの設計。一般社団法人として以下の実施機関から理事または評議員が出て実施委員会を構成し、各協力機関との連携した実施機関として責任を果たせる体制ができています。

東北大学、東京大学、京都大学、九州大学：医療データ流通共通プラットフォーム活用センター設置運用、関連システム開発

東京大学、大阪大学、京都大学、九州大学：臨床課題指向型次世代電子カルテシステムおよび標準的PHRの設計と開発

名古屋大学、帝京大学、九州大学：医療データ流通共通プラットフォーム活用センター運用実証試験

### ⑤ 所要経費

データベース構築を最初の3年、システム開発を最初の5年で行い、4年目以降7年間で運用とデータベースの維持拡張を行う。

総額：100億円（期間10年間）

1) 医学医療統合データ解析活用センター（東大）整備 総額45億円

初期経費（当初3年計）：臨床情報統合データベース構築 34億円 \*サテライト3箇所の経費を含む。

（内訳）健康医療情報収集・名寄せ・匿名加工化基盤 30億円（100施設分）、センター 4億円（データベース構築、開発）

運用経費（4年目から10年目計）：11億円（内訳）システム運営 0.5億円/年×7年 計3.5億円、センター人件費：1億円/年（15人）×7年 計7億円、その他：0.5億円

2) 開発研究 総額45億円

（内訳）統合データベース・解析システム研究開発：期間5年

計15億円、臨床課題指向型構造化次世代電子カルテシステム研究開発：期間5年 計15億円、PHR標準化と普及推進5年 計5億円、医用人工知能応用システム研究開発：期間5年 計5億円、データ解析と社会実装：期間7年 計5億円

3) 管理的経費 総額10億円（期間10年）

### ⑥ 年次計画

最初の3～5年でシステム設計、開発し、その後7年運用・拡張する。

1年目～5年目：臨床課題指向型構造化次世代電子カルテシステムの設計に関する研究開発を学会、工業会、関連団体等と実施する。2年次からは実際にシステム開発を開始し、全国の主要医療機関の既存情報システムとのデータ連動機構を各メーカーと共同研究開発し、主要医療機関に導入する。5年次から本格展開を開始し、FHDB集積を開始する。

FHDBの構築を認定匿名加工医療情報作成事業者と既存の国内の研究事業のリーダーの合同チームで行う。3～4年次はFHDBの開発と臨床課題指向型構造化次世代電子カルテシステム、生体センサーおよびIoTを駆使した標準化PHRとの連動設計を行う。FHDBを活用する医用人工知能応用システムの設計開発を行う。

「医学医療統合データ解析活用センターを」データベース拠点（東京大学を想定）に設置する。5年次から同センター機関の研究者による試験的データ利用サービスを開始する。全国公募により3箇所にサテライトを設置し、同センターに安全にネットワーク接続してデータ利用ができる環境を整備する。

6年目～10年目：FHDB等の実データを6年次から一般の研究者にも公開する。個々の研究者の多様な着想に基づき、研究デザイン構築、DBからのデータセット抽出とクレンジング、統計解析、論文執筆までのすべてのプロセスを支援する。一方、データ品質バリデーションを行い、必要な精度情報を医療機関側システムにフィードバックして精度修正を促すシステムの開発と運用を目指す。6年次からは全国3箇所のサテライトで一般の研究者がオンサイトで直接操作し解析できるサービスを運用する。FHDB等のデータを国内外の研究者が効率的に解析できる多次元で高度な解析ツールを開発し、同センターで提供する。

### ⑦ 社会的価値

医療ビッグデータは、様々な疾患や健康状態に関する悉皆データの収集・解析を実現できる。どの疾患がどこにどの規模で蔓延しているかという疫学情報、有効な診断・治療技術がどの医療機関で利用可能かといった患者に役立つ情報、個人のライフスタイルに即した健康情報などを、国民に直接提供できる。臨床アウトカム評価や薬剤副作用発生状況の把握、予期しない臨床イベント発生の解析などが、リアルタイムに低コストで実現できる。創薬や臨床研究の萌芽をもたらす。国際治験・臨床研究の基盤、他国との患者情報交換の基盤になり、経済的・産業的価値が大きい。また、保健・医療・介護に関する社会保障システム改革に資する知見を提供し、エビデンスに基づく医療経済政策の推進を可能にする。医療ビッグデータと医用人工知能応用は、世界で最早の超少子高齢社会を迎えた日本がこれを乗り越えるために不可欠なインフラ・ストラクチャーである。また医療ビッグデータの高度利活用は、世界の模範となるべき日本の保健・医療・介護システムのたゆまぬ改革に必須である。

### ⑧ 本計画に関する連絡先

中島 直樹（九州大学病院メディカル・インフォメーションセンター）

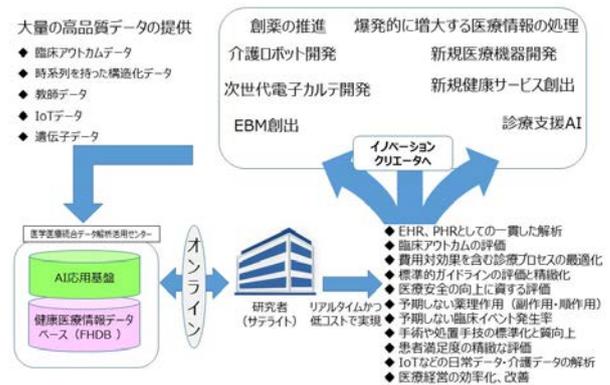


図2 研究成果と社会還元