

平成28年度研究開発成果概要書

採択番号：178D01

課題名：ソーシャル・ビッグデータ利活用・基盤技術の研究開発（平成28年度）

個別課題名：課題D 地域・社会課題解決のための異分野ソーシャル・ビッグデータの横断的利活用による近未来予測技術に関する研究開発

副題：行政が持つ健康情報と医療情報のデータベース構築と解析研究開発

(1) 研究開発の目的

医療系および健康系の複数のデータベースの情報を接続し、人生における健康史ともいえるライフコースデータを完成し、母子保健や学校健診と医療情報を接続する。これにより、どのように病気になっていくのかという解析による予防医療、発症予測や難病創薬へと応用する。また、どのような医療を受けた個人がその後高齢者時点でどのような健康状態となっているかという解析により、医療や薬剤の評価や費用対効果、介護における評価指標等、現在医療の様々な問題に対してソリューションを提案する基盤をつくる。

(2) 研究開発期間

平成28年度から平成30年度（3年間）

(3) 実施機関

国立大学法人京都大学（実施責任者 教授 川上浩司）＜代表研究者＞

(4) 研究開発予算（契約額）

総額60百万円（平成28年度20百万円）

※百万円未満切り上げ

(5) 研究開発項目と担当

研究項目1：学校健診情報データの収集のための自治体との連携交渉とデータ収集

1. 自治体との連携交渉
2. 連携した自治体の中学校へデータ収集

研究項目2：匿名化してスキャンした情報からデータベース構築まで

1. データ化に際する個人情報部分の暗号化
2. 健康情報レポート還元システム

研究項目3：母子保健情報のデータベース化と学校健診情報と合わせての解析

1. 母子保健情報のデータ収集
2. 母子保健情報データベース構築
3. 疫学研究による疾病予測方法の開発

研究項目4：電子カルテ情報の収集による診療情報データベース構築と予測モデルの検討

1. 診療情報データベースの構築
2. 糖尿病治療経過から重症化を予測する研究

(6) これまで得られた成果（特許出願や論文発表等）

		累計（件）	当該年度（件）
特許出願	国内出願	0	0
	外国出願	0	0
外部発表	研究論文	1	1
	その他研究発表	15	15
	プレスリリース・報道	7	7
	展示会	0	0
	標準化提案	0	0

(7) 具体的な実施内容と成果

研究項目1：学校健診情報データの収集のための自治体との連携交渉とデータ収集

1. 自治体との連携交渉

平成 28 年度までに 30 自治体以上を目標として、全国自治体と情報提供や交渉活動を進めた。現在 34 都道府県 86 自治団体と連携交渉を調整しており、平成 28 年度事業として、43 自治体+私立 2 校、24,579 人分のデータ化、データ取得を完了済である。

2. 連携した自治体の中学校へデータ収集

項目 1-1 で連携を契約した自治体の中学校へ出向き、スキャン作業を行う。当初の目標とした H28 年度末までに 30 自治体以上、という目標を達成し、本年の事業は 43 自治体からのデータ収集を完了した。

研究項目 2：匿名化してスキャンした情報からデータベース構築まで

1. データ化に際する個人情報部分の暗号化

本課題の開始前には調査票のスキャン開始時に帳票の匿名部位をあらかじめ指定しなければならず、フォーマットが異なる場合などでは特に時間を要していた。未知のフォーマットであっても、自動での匿名部位の特定ができるよう研究開発を行うことを目標とし、スキャン時に自動で匿名部位の特定ができるマスキング技術を開発した。

2. 健康情報レポート還元システム

個人向けレポートにおいては、文部科学省や学校現場からの要請を踏まえて成長曲線の自動記載などの新機能を開発すること、また、自治体向けレポートにおいては、自治体内地域や学校毎の比較をより精緻にできるためのレポート自動化を目標とした。成長の記録としてレポートを提供し始めている。

研究項目 3：母子保健情報のデータベース化と学校健診情報と合わせての解析

1. 母子保健情報のデータ収集

交渉によるデータ提供の合意取得とデータの収集を目標として進め、現在、10 以上の自治体からデータ提供合意、5 自治体よりデータ取得している。一方、乳幼児健診は全国で悉皆での実施とは言え様式が統一されていない点が課題である。

2. 母子保健情報データベース構築

データベース構築にあたり、各自治体からの多様なデータ構造を受入れる一次データベースを構築することを目標としている。標準的なデータ構造を構築した上で、疫学研究に利用可能な二次データベースを構築すべく取組みを進めている。

3. 疫学研究による疾病予測方法の開発

母子保健情報と学校健診情報を活用して、各種の疫学研究を行うための準備を進めている。防府市、津野町の母子一学校情報を試験的に接続した。成長過程や疾患への罹患を対象として研究を行い、疾病予測方法の開発に寄与する。

研究項目 4：電子カルテ情報の収集による診療情報データベース構築と予測モデルの検討

1. 診療情報データベースの構築

電子カルテのベンダーや形式に寄らず個人情報を削除、匿名化した形で情報収集し、データベースを構築することを目標に、日本全国の民間病院、大規模診療所や公立病院との交渉を行い、電子カルテ及びレセプト情報の収集に努めている。本研究期間中（～30 年度）の目標を 100 医療機関、500 万人規模のデータベースとしていたが、現在全国約 130 の医療機関と契約し、契約ベース 1400 万人、うち 450 万人のデータが格納済であり、事業は極めて順調に進捗している。

2. 糖尿病治療経過から重症化を予測する研究

電子カルテから構築した診療情報データベースを用いることで、糖尿病の治療経過と合併症・重症化を予測することが可能になると考えられ、これを目標に研究を進めている。