

(29-2)

様式1-4-2

平成 29 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

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に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)

(4) 研究開発予算 (契約額)

総額 60 百万円 (平成 29 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

研究開発項目 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

1-1 自治体との連携交渉

全国自治体と、さらなる情報提供や交渉活動を進める。現在 41 都道府県 113 自治体と連携を調整しており (図 1 参照)、平成 29 年度、提出時点の取り組み参加自治体数の実績は 70 自治体と順調に事業拡大を進めている。

1-2 連携した自治体の中学校への説明とデータ収集

取り組み参加自治体において、各中学校に出向き、健診票のスキャン作業を行なった。

研究開発項目 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

2-1-1 データ化に際する個人識別情報のマスク

平成 28 年度までの実績として、学校健診情報において、スキャン時に自動で匿名部位の特定ができるマスクング技術を開発・実装した。平成 29 年度は、その技術を、母子保健における乳幼児健診表の紙媒体のスキャン時にも応用できるように開発を行なった。

2-1-2 個人や自治体向けの健康情報レポート還元システム

平成 28 年度では参加自治体内の中学校の生徒に対して、健康成長の記録としての健康レポートを還元するシステムを開発した。平成 29 年度では、そのシステムを洗練し、自治体数が増加しても迅速に処理できるように改修し、かつ、レポート内容の項

目追加、改善を行なった。

研究開発項目 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

3-1 母子保健情報(母子保健法に基づく乳幼児健診)のデータ収集

平成 29 年度、提出時点で、母子保健情報は、既に下記 11 自治体からデータ収集を完了している。

3-2 母子保健情報データベース構築

平成 29 年度は、母子保健情報の収集・蓄積システムを構築し、試験的に健診項目を限定して、試験的に運用し、DB 格納、レポート還元を 4 自治体の実績として行なった。また、健診項目を格納するための標準化マスタの開発を行なった。

3-3 母子保健一学校健診を活用した疫学研究による疾病予測方法の開発

平成 29 年度は 6 自治体を対象として、母子保健情報と学校健診情報の接続を行なった。生年月日と名前にもとづいて母子保健情報と学校健診情報を接続するシステム開発を行い、実装し、乳幼児期～学童期までをつなげたデータセットを作成し、その解析に着手した。

研究開発項目 4：糖尿病治療経過から重症化を予測するための研究

4-1 電子カルテ情報の収集による診療情報データベースの構築

平成 29 年度はデータ抽出可能なベンダーの増加に注力を注ぎ、医療機関数を増やすこととした。また、検査・薬剤・病名データの標準化作業も並行して行った。その他、レセプト DPC と電子カルテとの医療機関内での突合の仕組みの開発も行った。

4-2 患者転帰を予想するアルゴリズムの妥当性を検証する際のデータソースとして電子カルテ情報の有用性が着目されている 1。今回は日本人の糖尿病患者の重症化予測モデル 2 の妥当性を、糖尿病患者 7-8 万人を含む電子カルテ情報で検証する。具体的には京都大学医の倫理委員会への申請・承認、データの取得、データの解析を経て論文文化を行う。

1: BMJ 2016 353:i3140、2: Diabetes Care 2013;36:1193

(6) 特許出願、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 | 16 | 15 |
| | その他研究発表 | 31 | 16 |
| | プレスリリース・報道 | 10 | 3 |
| | 展示会 | 0 | 0 |
| | 標準化提案 | 0 | 0 |

(7) 具体的な実施内容と成果

研究開発項目 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

1-1 自治体との連携交渉

全国自治体と、さらなる情報提供や交渉活動を進める。現在 41 都道府県 113 自治体と連携を調整しており (図 1 参照)、平成 29 年度、提出時点の取り組み参加自治体数の実績は 70 自治体と順調に事業拡大を進めている。

1-2 連携した自治体の中学校への説明とデータ収集

取り組み参加自治体において、各中学校に出向き、健診票のスキャン作業を行なった。

研究開発項目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

2-1-1 データ化に際する個人識別情報のマスク

平成28年度までの実績として、学校健診情報において、スキャン時に自動で匿名部位の特定ができるマスクング技術を開発・実装した。平成29年度は、その技術を、母子保健における乳幼児健診表の紙媒体のスキャン時にも応用できるように開発を行なった。

研究開発項目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

3-1 母子保健情報(母子保健法に基づく乳幼児健診)のデータ収集

平成29年度、提出時点で、母子保健情報は、既に下記11自治体からデータ収集を完了している。

3-2 母子保健情報データベース構築

平成29年度は、母子保健情報の収集・蓄積システムを構築し、試験的に健診項目を限定して、試験的に運用し、DB格納、レポート還元を4自治体の実績として行なった。また、健診項目を格納するための標準化マスタの開発を行なった。

3-3 母子保健—学校健診を活用した疫学研究による疾病予測方法の開発

平成29年度は6自治体を対象として、母子保健情報と学校健診情報の接続を行なった。生年月日と名前にもとづいて母子保健情報と学校健診情報を接続するシステム開発を行い、実装し、乳幼児期～学童期までをつなげたデータセットを作成し、その解析に着手した。

研究開発項目4 糖尿病治療経過から重症化を予測するための研究

4-1 電子カルテ情報の収集による診療情報データベースの構築

現在12ベンダーの電子カルテから、個人情報の匿名化と収集を可能にするシステムを構築しているが、引き続きあらゆるベンダーに対応できるよう開発を継続していく。データを提供いただく医療機関についても、現在101医療機関まで増えているが、引き続き交渉を行い、日本を代表するような医療データベースが構築できるよう、引き続き交渉を継続する。

4-2

日本人の糖尿病患者の重症化予測モデルの妥当性を全国の糖尿病患者7-8万人を含む電子カルテ情報で検証予定であったが、入手が見込まれるデータセットに喫煙情報など予測モデルに必要な変数が不足していることが判明した。腎症の進展抑制が社会的問題となっているため、新規開発されたSGLT2阻害薬が他の血糖降下薬に比べ腎機能抑制効果を有するかを検討する。研究テーマについての京都大学医の倫理委員会への申請・承認を得、今年度内にデータセット入手する。次年度解析・論文化を行う予定。

(8) 外国の実施機関