

地球一括計上

課題名	環境因子の変化に伴う疾病構造変化モニタリングと中長期環境モニタリング及び曝露調査結果を用いた環境がヒトへ与える影響の解析を行う病院コホートを利用したデータマイニングシステムの研究		
担当研究機関	国立研究開発法人国立国際医療研究センター		
研究期間	平成24－28年度	合計予算額 (当初予算額 ベース)	93,266千円（うち28年度 14,265千円）
研究体制	国立研究開発法人国立国際医療研究センター		
研究概要	<p>1. 序（研究背景等）</p> <p>本研究の背景として、公害病についての研究が主な環境と疾病との問題点であった前世紀に比較し、現代においては温暖化をはじめとするグローバルな気候変動、PM2.5等の大気汚染物質に関する社会的な認知度が向上し、これら環境因子の健康影響に対する関心も上昇している。公害に対する理解から、各国は環境基準を策定している。適切な環境基準に基づき、環境行政・民間が共働して環境を維持することで、多くの先進国において重度の公害病は減少している。しかしながら、環境因子の変化に伴い人が罹患する疾病は変化していく。温暖化をはじめとし、地球規模での環境変化による健康影響を調査するための基盤づくりに、各国では環境因子のモニタリングがさかんに実施され、いままで環境因子の影響が指摘されてこなかった、あるいは経験的に指摘されていても、エビデンスが得られていなかった疾病に関して、環境因子と疾病構造に関する研究が萌芽的に始まっている。我が国においても、全国を網羅し、一般公開されている環境ビッグデータが存在し、詳細な気象データやPM2.5を代表とする新たな環境因子が精密にモニタリングされている。そして徐々に新たな環境因子が疾病構造に影響を及ぼしていることが知られつつある。これら、従来研究の対象よりはるかに微細な環境因子の変動が、いかにわれわれの疾病構造に大きな影響をあたえているかを解析し、新たな知見を得るためには、いままで以上により大規模な症例群から、より詳細な臨床データを集めた、大きなデータベースが不可欠である。しかし、現在環境モニタリング結果と診療情報システムデータとを関連付けたデータ解析を行った事例は近年でもいまだに少数である。これは、最近まで診療情報が電子化されてこなかったこと、電子化されたデータへのフルアクセスが医師のみに制限されがちであること、診療データの記録目的が診療のためであり研究目的ではないことからデータに2次利用を行うためには様々な前処理、ノイズ処理が必要となることが障壁となってきたためである。また、研究の実現に情報科学の専門家と医師・医学者の協力が必要不可欠となるが両者の連携は専門性の違いから今までは行われることが少なかったなど、様々な技術的要因・社会的要因から環境モニタリング結果と診療情報の突合解析が行われてこなかったためである。</p> <p>2. 研究目的</p> <p>本研究では、情報科学と医学の集学的な取り組みにおいて、環境データと診療情報とを関連付けたデータ解析を行い、今後複数の協力施設と連携してデータ解析を行うことを念頭に、情報科学的な手法を駆使して、効率よくかつ個人情報に配慮して安全に、低コストで迅速な突合解析システムを構築するため、具体的な解析を通じて、環境因子との関連性が強い疾病候補を絞り込み、継続的な環境因子の変動と疾病構造の変化のモニタリングを可能にするインフラを構築することを目的としている。この事例として、病院情報システムのデータと、環境省が保有する環境データ（環境省大気汚染物質広域監視システムのデータ等）の統合・解析を可能にする一連のシステムや解析過程を統合、一部可能な部分を自動化した「共通システム」を構築する。構築した「共通システム」は、システム本体の対価をもとめず、本研究の目的である、今後複数の協力施設と連携してのデータ解析を実現するためのインフラ整備として、学術誌などを用いて一般に公開し、導入希望の医療施設等をつのり、希望施設があればシステムの提供を目指していく。</p>		

3. 研究の内容・成果

今年度は研究の最終年度であるため、本報告では下記3点について述べる。

- I. 本研究を通じて構築した「共通システム」の全貌を解説する。
- II. 本研究班が行った具体的な応用解析例について紹介する。
- III. 今後複数の協力施設と連携してデータ解析を行うために「共通システム」を無償配布する目的で行う、学術的に公表・導入可能な施設への呼びかけについて説明する。

I. 本研究を通じて構築した「共通システム」について

病院情報システムと環境ビッグデータ突合解析のための「共通システム」の概要として下記のアルゴリズムで解析を行うシステムの総称である。

① 病院情報システムビッグデータからの「臨床データベース」構築

「臨床データベース」の蓄積成果

国立国際医療研究センター

2011年～2016年 自動解析用処理済みデータセット

匿名化患者 572384件

匿名化処方 1290253件

匿名化病名 239112件

匿名化予約 608147件

五百山クリニック 2014年

非アレルギー性鼻炎（NAR）の患者データ

著名改善 130件

改善 88件

不変・中断 45件

悪化 1件

② 環境ビッグデータからの「環境データベース」構築

「環境データベース」の蓄積成果

- ・抽出対象地区 1,749件
- ・時間単位データ 15,372,000件 (92,064,000件2011年～2016年累積値)
- ・日単位データ 640,134件 (3,833,808件2011年～2016年累積値)

③ 「臨床データ・環境データ突合済みデータベース」

④ 未知の知見候補抽出へのスクリーニング解析

平成27年までの成果により機能している「臨床データ・環境データ突合済みデータベース」を利用し、今年度の成果として、解析対象疾病候補が未定の場合を想定し、ある特定の環境因子に関連して疾病構造が変化する疾病候補を抽出するための、未知の知見候補抽出へのスクリーニング解析を行うシステムを開発・自動化した。

(イ) 「四分位オッズ解析システム」（平成28年度自動化）

(ロ) 臨床疫学的な解析候補の絞り込み作業工程（平成28年度ソフトウェアとして統合）

⑤ データベースをコホート研究に利活用する応用研究手法

(イ) 特定の疾病に関する複数の環境因子の影響解析プロセス

(ロ) 「オッズ比を用いた混合モデルによる電子カルテデータの病院コホート研究システム（平成28年度自動化）

⑥ 「共通システム」（①から⑤で構成される一連の解析システムの総称）から「環境モニタリングシステム」へ

共通システムを中長期に継続的に運用し、今後、複数施設に導入・連携していくことで、多岐にわたって応用可能な、疾病構造解析に関わる「環境モニタリングシステム」として活用することができる。図1に、本研究を通じて構築した、病院情報システムと環境ビッグデータ突合解析のための「共通システム」とその解析アルゴリズムをフローチャートで示す。

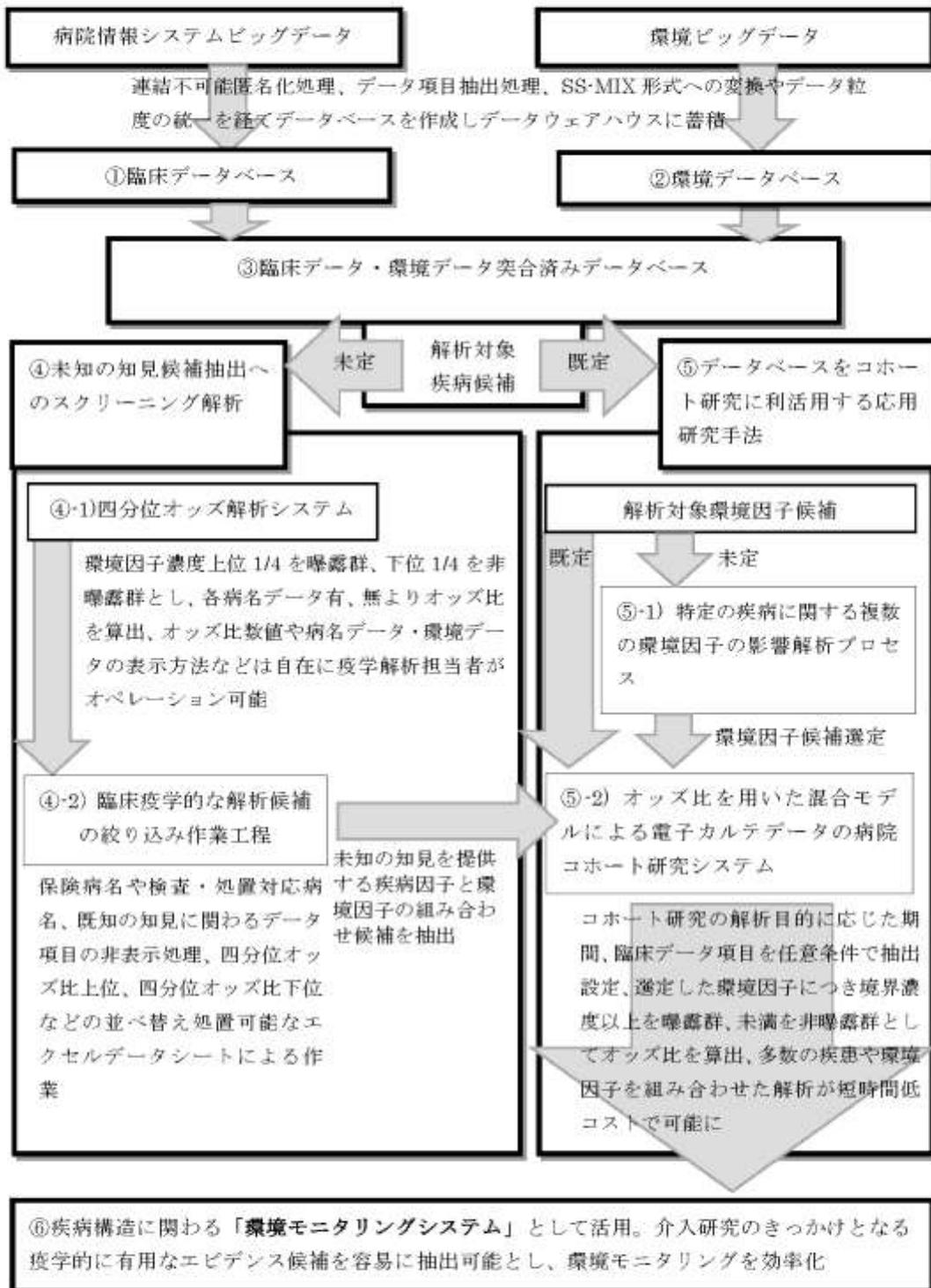


図 1 病院情報システムと環境ビッグデータ突合解析のための「共通システム」の概要

II. 本研究班が行った具体的な応用解析例

情報学的手法を駆使して、いままでの紙カルテベースでは収集不可能であった、膨大な症例データを電子カルテビッグデータから抽出、解析する方法が、今年度完成した「共通システム」である。本研究では、環境因子データと、疾病情報を含んだ電子カルテビッグデータを突合解析し、解析過程を含め学術誌に発表することで、この方法論で得た疫学データを社会に還元してきたが、今年度新知見として解析が現在進行中の応用解析成果の実例には下記がある。

- ① 「共通システム」を用いて抽出された解析対象候補疾患の抽出
- ② 喘息患者群の夏日における PM2.5 曝露と嚥下性気管支炎・肺炎の発症との関連性分析
- ③ 寒冷環境下での PM2.5 曝露と糖尿病高齢者における曝露 2 日目の虚血発作発症四分位オッズ比の PM2.5 濃度依存性の解析
- ④ 温暖時期の最高気温と糖尿病既往のない全年齢患者群におけるくも膜下出血発症との関連

III. 今後複数の協力施設と連携してデータ解析を行うために「共通システム」を無償配布する目的で行う、学術的なシステムの公表・導入可能な施設への呼びかけ

本年度の研究実施計画にも記載あるように、本研究で開発した「共通システム」は、病院情報システムのデータと、環境省が保有する環境データ（環境省大気汚染物質広域監視システムのデータ等）の統合及び解析を可能にするシステムであり、今後、環境因子が疾病構造に与える影響を継続的にモニタリングし得る基盤となるシステムである。このため、本年度の成果として、当該システムを公開し、照会に応じて配付する日本語版マニュアル一式等を完成し、複数の協力施設と連携してデータ解析を行い得る環境の整備を図った。具体的な公開方法として、下記2通りの方法を用いた。

(ア) 研究会における公表

本研究班が主体となっている、電子カルテデータ・環境データを統合した疾病モニタリングに関する研究会主催の講演会において、参加者に「共通システム」を公表した（「医療分野におけるオープンソースソフトウェアについての現状と最近の問題点について、GNU Healthの動向を含めて」主催；電子カルテデータ・環境データを統合した疾病モニタリングに関する研究会，於 国立国際医療研究センター，講演：京都大学 医学研究科／EHR共同研究講座／特定講師 小林慎治，平成28年11月18日）。

(イ) 医療分野におけるオープンソースソフトウェアについての現状および病院コホートを 利用した医療情報と外部データ突合用データマイニングシステムの紹介

さらに広く、同様の電子カルテシステムを擁して、地域の中核的な医療センターとなっている、国立病院機構の病院群に「共通システム」をアピールし、学術的な公開を行う目的で、国立病院機構の査読有の学術誌である「医療」にシステムの概要とソフトウェア部分の無償配布、導入希望施設を広く募る旨の呼びかけを明記した原著論文を投稿

4. 考察

共通システムを中長期に継続的に運用し、今後、複数施設に導入・連携していくことで、多岐にわたって応用可能な、疾病構造解析に関わる「**環境モニタリングシステム**」として活用することができる。また、介入研究のきっかけとなる疫学的に有用なエビデンス候補が容易に抽出可能となり、環境モニタリングとその応用が飛躍的に効率化され、イノベーションまでのスピードアップが期待される。実際、本研究班でも先述した「オッズ比を用いた混合モデルによる電子カルテデータの病院コホート研究システム（平成28年度自動化）」開発にむけて、実装しながら改良を繰り返すスパイラルモデルによる開発の過程で、解析候補抽出のためのプレスタディを積み重ねながら、II部で述べる多数の解析結果を得た。このような解析結果を生気象学者や臨床疫学者、統計学の専門家が集学的に検討することで、介入研究のきっかけとなる疫学的に有用なエビデンス候補抽出が効率的に抽出、必要に応じて中長期のモニタリング候補に設定することも可能であるため、将来的なモバイル端末を介したリアルタイムの健康天気予報など、イノベーションに貢献する基盤となると考える。

5. 波及効果

本報告書作成時点では特になし。