

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）

平成30年度 統括研究報告書

周産期関連の医療データベースのリンケージの研究

研究代表者 森崎菜穂 国立成育医療研究センター社会医学研究部・室長
(森臨太郎 国立成育医療研究センター政策科学研究部・部長)

研究分担者 森臨太郎 国立成育医療研究センター政策科学研究部・部長
掛江直子 国立成育医療研究センター生命倫理研究室・室長
康永秀生 東京大学大学院医学系研究科公共健康医学専攻臨床疫学・教授
大田えりか 聖路加国際大学大学院看護学研究科・教授
永田知映 国立成育医療研究センター臨床研究開発センター
臨床研究教育部・室長
森崎菜穂 国立成育医療研究センター社会医学研究部・室長
Mahbub Latif 聖路加国際大学大学院 公衆衛生学研究科

研究要旨

本研究は、周産期に関連する各種データベースとの連結可能性を試行することで将来の有効的な活用を促すこと、また、他データベースとの連結を通して各データベースの妥当性を測ることを目的とした。

研究3年目である本年度は、i) 各分担の先生方と産婦人科医・小児科医・疫学者の協調を促し、周産期に関連する各種のデータベースを連結したデータベースの解析を行い単一のデータベースからは産出不可能であったエビデンスを産出する、□)他の大規模データベースを連結するための整備を行う、ことを目的として、これを実践した。

A. 研究目的

本研究では成育医療分野における各種統計や医学団体所有データベースを過去に活用し、また周産期医療関係の各種データベ

ースをリンケージする手法に関する研究を行ってきたという経験を生かして、成育医療分野のデータベースを連結することで拡充し、さらに多くの臨床研究に活用する。

また、公的統計の妥当性検証やデータベース同士の自動連結手法を確立することで今後の研究基盤を作成することが目的である。

これにより、医療計画の「5 疾病 5 事業」に含まれ、その医療体制整備は極めて重要である小児医療および周産期医療において、関連学会が積極的に作成してきたレジストリや政府が行ってきた政府統計をいまよりも更に有用に活用する方法が広がると考えられる。

本研究での今年度の主な研究目的は下記の通りであった。

-日本語に対応した自然言語解析の技術の一つである InterSystems IRIS Natural Language Processing (NLP) Japanese を利用し、レセプト傷病名の自動類型化が可能であるかを検証する。

-周産期医療の入院診療の処置・投薬についての情報をレセプトから抽出し、新生児臨床データベース(NRN)などの臨床レジストリを代替するシステムを作れるかを検証する。

-平成 29 年度から引き続き、人口動態統計の出生票および死産票に記載されている児の母の情報と、女性の死亡票を高精度にリンケージすることで、妊娠後の女性の死亡を把握するための方法を検証する。

-平成 28 年度から周産期関連の全国データベースや、これら同士を連結することで得られたデータベースを多角的に解析することで、妊婦および出生児の長期予後について、成育医療分野に有用なエビデンスを

提供する。

-既存の医療情報を医学研究のために利活用するに際して、適用される現行法令や指針にはどのようなものがあり、また、各法令等にはいかなる事項等が規定されているのかにつきまとめる。

B. 研究方法

本研究は、複数のデータベースを相互に利活用しながら研究を行っていくという性質上、各分担研究班の分担研究者および研究協力者同士がお互いの研究を補助するという緻密な連携を取りながら行った。また、本研究を実施するにあたり、医療データベースのリンケージに関する倫理的・法的側面の妥当性についても、研究分担者である倫理専門家の監督のもとで行った。

なお、「小児死因統計の臨床的死因との相致性に影響する要因に関する研究」については、別研究（厚生労働科学研究費補助金（健やか次世代育成総合研究事業）

「突然の説明困難な小児死亡事例に関する登録・検証システムの確立に向けた実現可能性に関する研究」(主任研究者 溝口史剛))にて継続実施されることになったため、本年度は本分担研究では行われなかった。

また、研究代表者の職場異動に伴い、年度途中に研究代表者の交代があった。

C. 研究結果

各種厚生労働省統計と周産期関連学会データベースのリンケージと解析

(分担：森崎菜穂、康永秀生、Mahbub

Latif)

森崎分担研究では、人口動態統計の出生票、死産票、および死亡票をリンケージする複数の手法を比較検討することで、もっとも正確にこれらをリンケージできる手法を提案すること、各種の周産期関連データベースをリンケージしたデータベースの利用を促進し、その解析を通して単一のデータベースからは産出不可能であった医学的なエビデンスを複数提示すること、を目的としている。

また、康永分担研究では、DPC 情報の他データベースがリンケージに関するフィージビリティ評価や、他データベースによる DPC 情報から算出された項目の妥当性評価を行うことを目的としている。

3年目である本年度は、2つの分担研究班と新規の分担研究者（生物統計家）が協働することで、2014-2016年度に出生した児の出生票と母の死亡票を高精度にリンケージするための方法を検討し、2つの分担班が協力することで、DPC 情報による周産期医療に関するレジストリ情報の代替可能性について検討し、日本産科婦人科学会周産期登録データベース、新生児医療ネットワーク登録データベース、出生票、死産票、乳児死亡票、および DPC 情報を様々な角度から解析し、妊婦および児の予後に関係する医学的・社会的因子について、複数のエビデンスを発表した。

小児慢性特定疾患データベースのリンケージと解析に関する研究

（分担：森臨太郎）

本分担班では、各分担班同士の情報共有を促し、各種データのリンケージ及び利活用を推進し、各種団体が所有するデータベースのリンケージを可能とするための整備を行っている。特に他データベースとのリンケージの実現可能性が高い二つの年間10万件を超える登録のある世界最大級の小児疾病登録データベース、そして早産児の長期予後に影響を与える社会的および医学的因子を把握するために、現在遂行中の早産児コホート（INTACT データベース）について、他データベースとリンケージすることの実現可能性を評価している。

本年度は、日本語に対応した自然言語解析の技術の一つである InterSystems IRIS Natural Language Processing (NLP) Japanese を利用し、レセプト傷病名の自動類型化が可能であるかを検証した。IRIS NLP Japanese は主たる意味を持つ語句を認識し、さらに語句の欠損や入替があっても標準病名との対応を自動的に行えることが分かった。IRIS NLP Japanese を利用した傷病名の類型化は、現実的な解法との一つとなり得る可能性があることが分かった。

周産期臨床データベースと DPC データを用いた、産科合併症に関する研究

（分担：永田知映）

本年度は、昨年度に引き続き調査票情報の提供を受けた人口動態調査（出生票・死亡票・死産票）を用いて、生殖可能年齢の女性の死亡票と、出生票・死産票をリンク

することで、出産あるいは死産から一定期間内に起こった死亡を網羅し、妊産婦死亡統計データの信頼性および母体死因に関する検討を行った。平成30年度は、電子化されていない個票データについても電子化を行い、リンケージ方法を改善して解析を行った。2015年、2016年の生殖可能年齢の女性の死亡票・死亡個票のリンケージにより死亡データベースを作成し、その死亡データベースと、2014年から2016年の出生票・出生個票のリンケージによる出生データベースと死産票・死産個票のリンケージによる死産データベースをさらにリンケージして、妊娠中から産後1年未満の女性の死亡357例を抽出した。そのうち最も多かった死因は自殺の102例であった。

死亡データベースに出生・死産のデータベースをリンケージすることで、死亡した女性の出産に関する情報を把握することが可能となり、産褥婦の自殺例や後期妊産婦死亡例の把握に有用であった。

また、現在の妊産婦死亡統計データおよび日本産婦人科医学会妊産婦死亡症例検討委員会のデータベースと比較検討することにより、データ間での解離の有無と妊産婦死亡に関連する因子を検討するために、日本産婦人科医学会の担当部署との調節を行った。

産褥婦の自殺にかかる状況及び社会的背景に関する研究

(分担：大田えりか)

本分担研究では、人口動態統計出生票及び

死亡票の連結により抽出された、2015～2016年における生児出産後1年未満の褥婦の自殺例92件を抽出し、背景や自殺方法などを分析した。35歳以上、初産、及び世帯の職業が無職の女性において、最も自殺率が高かった。自殺の時期は、産後1年を通して自殺がみられた。人口動態調査出生票及び死亡票のリンケージは、産褥婦の自殺死亡例の把握に有用な手段と考えられる一方で、死亡診断書に記載される事項は限られているため、産褥婦の自殺死亡を予防するための対策に結びつけるためには情報が不十分であり、各症例についてさらなる詳細な調査が必要と考えられる。

医療データベースのリンケージ研究実施に際して適用される法令等に関する検討

(分担：掛江直子)

本分担研究では、既存の医療情報を医学研究のために利活用するに際して、適用される現行法令や指針にはどのようなものがあり、また、各法令等にはいかなる事項等が規定されているのかにつき検討を行った。

個人情報の取扱いについて適用される法令は、各医療情報を保有する機関によって異なり、また、各法令の規定が一律ではない。利用目的、情報の内容、情報を保有する者や情報を取扱う主体によって適用法令や該当条文は異なり、また、情報の内容については、法の規制対象外の情報、法令の規制を受ける情報であっても、どのような質(種類)の情報であるか、すなわち、個人情報、

要配慮個人情報、対応表により本人を識別することができる匿名化された個人情報、本人を識別することが不可能な匿名化された個人情報、統計情報であるか等によって、情報の取扱い等にかかる適用法令や該当条文が異なっており、遵守事項に違いが生じる。また、法令が規定する個人情報の取扱いの他に、日本の研究機関で実施され、または日本国内で実施される人を対象とする医学系研究については、医学系指針の遵守が必要となるが、当該指針の適用除外となる人を対象とする医学系研究もある。

個人情報たる医療情報の取扱いがこのような複雑な法制度となっていることにつき、運用・遵守しやすいシンプルなルールを作ることの必要性や、個人情報の取扱いのあり方や考慮・保護すべき事項の再検討を含む抜本的な対応が必要である等の指摘がなされている。

D. 考察

本年度は研究の最終年度として、研究 1-2 年目に活用してきた人口動態統計や周産期医療に関する臨床レジストリの継続的な利活用のみならず、レセプト・データベースやDPC情報へのリンケージの拡大を検討した。また、これらのデータとのリンケージ体制の構築と実務を通して、現状の可能性と問題点とまとめた。

また、引き続き、得られたデータベースの解析から、妊婦および児の予後に関する医学的・社会的因子について、産科医・小児科医・疫学者とともに複数のエビデンスを

発表した。

さらに、今後増加が見込まれる医療情報の二次利用による臨床研究を行う際に留意すべき現行法令や指針とまとめ、今後の参考資料とした。

「医療分野の研究開発に資するための匿名加工医療情報に関する法律」(平成 29 年 5 月 12 日公布)の成立を受けて、日本でも医療データを他のデータベースと連結し活用することが促進されると思われる。

本研究では、データベースをリンケージし解析するという実践を通して、数々の臨床・社会的課題にそぐう解析を行い実際多くの研究論文の刊行に成功することで、データ・リンケージにどのようなその可能性が潜んでいるのかを示すことができた。

現在、他分野のデータベースとの連結可能性についても模索が行われている。今後も、データ・リンケージにより質の高いエビデンスが産出できるようなシステム作りや、そのデータが活用できる分野の拡大が望ましいと思われる。

E. 結論

研究最終年度である本年度は、リンケージを行うデータベースの更なる拡張を行うと同時に、これらのデータとのリンケージ体制の構築と実務および医療情報の二次利用による臨床研究を行う際に留意すべき現行法令や指針とまとめることを通して、医療情報のデータ・リンケージとその利活用の可能性と問題点とまとめた。