

令和元年6月5日現在

機関番号：11301

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K20899

研究課題名（和文）妊娠高血圧症候群の高精度な予測と新規知識獲得のための人工知能の開発

研究課題名（英文）Building of work-flow to develop machine learning model for prediction of Hypertensive disorders of pregnancy (HDP)

研究代表者

水野 聖士 (SATOSHI, MIZUNO)

東北大学・東北メディカル・メガバンク機構・助手

研究者番号：80646795

交付決定額（研究期間全体）：（直接経費） 3,000,000円

研究成果の概要（和文）：本研究では、生理学検査や尿検査、既往歴などの多くの情報源を統合的に解析しHDPの病態分類を行う高精度なフェノタイピングアルゴリズムの開発と、得られた病態分類の結果を正解とし、機械学習モデルの検討を行い、大規模出生コホート調査でのHDPの早期診断支援のための機械学習モデル構築のためのワークフローを検討した。その結果、大規模出生コホート調査でのHDPの早期診断支援のための機械学習の可能性が示唆された。

研究成果の学術的意義や社会的意義

本研究で検討を行った大規模出生コホートでの妊娠高血圧症候群のフェノタイピングアルゴリズムによる病型分類と、病型分類の結果を正解ラベルとした機械学習モデル構築のワークフローは、三世代コホート調査における妊娠高血圧症候群の研究の重要な情報リソースの一つとなりうる。

研究成果の概要（英文）：We considered the work-flow to develop supervised machine learning model for prediction of hypertensive disorders of pregnancy (HDP) with large-scale birth cohort dataset. To obtain class-label of supervised learning, we developed rule-based phenotyping algorithm according to clinical guidelines. The developed algorithm was applied into phenotyping of Birthree cohort subjects. We tried to develop the ML model to predict HDP with phenotyped disease types as class labels of supervised machine learning. In this study, we used Birthree cohort data before data-cleaning to develop the work-flow. In the future work, we will study with data after cleaning to develop precise informatics bases of HDP study in Birthree cohort study.

研究分野：バイオインフォマティクス

キーワード：バイオインフォマティクス 機械学習 フェノタイピング 早期診断支援 深層学習 産科 妊娠高血圧症候群 HDP

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

(1) 糖尿病・ガンなどの多因子疾患についての多くの研究から、多因子の複合的な相互作用が疾患発症に大きな影響力をもっていることが知られるようになる一方で、多くの疾患について発症予測、発症機序の解明、有効な介入方法の開発が困難である状況が続いている。これらの状況を踏まえ、遺伝因子、曝露因子などの大量のデータを統合し、疾患の発症予測、治療法の確立、治療結果の最大化を行う精密医療の概念が提唱されているが、これを実現するためには、巨大なデータを読み解き、新たな知見を発見する深層学習モデルを備えた人工知能が求められる。深層学習は、これまでの方法で変数選択に推論の結果が強く縛られていた問題を解決し、正確な推論を多くの問題に対して適用可能となることが期待される。

(2) 妊娠高血圧症候群(HDP)は、全妊婦の約7%に起こり、全妊婦の1~2%が重症患者である高頻度かつ重症化しやすい疾患である。重症HDPは母体の周産期死亡のリスクを高めると共に、胎児機能不全や早産を誘発し、妊娠中断や児の後遺症リスクを増大させるため、児の生命や予後への影響も大きい。これまでに、HDPの原因解明のために疫学研究、家族性連鎖研究などが行われ、遺伝因子・曝露因子を含む多くのリスク因子の関与が指摘されると共に、発症予測のために多くの研究がされているが、その病態は不明であり、有効な発症予測が可能な変数の組み合わせは見つかっていない。

(3) これまでの研究で学習モデルの解析結果から新たな知識の獲得を行うための、HDPに関する網羅的な非構造化知識の取り扱いに必要な、Pre-eclampsia Ontology(PEO)を構築している。また、予備研究として、母集団におけるHDPの発生率や変数の分布の偏りを反映し人工的に生成したデータを使い深層学習を使用した高性能な学習モデルの開発に成功している。そこでこれまでの研究を拡大し、妊娠高血圧症候群の高精度な予測および、予測結果から新規知識獲得を行うための機械学習モデルの開発を計画した。

2. 研究の目的

本研究では、HDPの予測のための機械学習の構築のためのワークフローの検討のため、以下の2つの目標を設定した。

(1) 医師の診断病名だけでなく、生理学検査や尿検査、既往歴などの多くの情報源を統合的に解析しHDPの病態分類を行うフェノタイピングの、大規模出生コホートへの適用可能性を検討する。

(2) 機械学習モデルの構築のための、大規模データの適切な前処理方法の検討と、学習モデル構築方法の検討を行い、大規模出生コホートにおけるHDPの早期診断支援のための機械学習の可能性を検討する。

3. 研究の方法

(1) HDPのフェノタイピングアルゴリズムの開発

国際的なHDPの診断基準に従うルールベースのフェノタイピングアルゴリズムの開発を行った。開発したアルゴリズムを、東北メディカル・メガバンク計画三世代コホート調査の対象者に適用するとともに、医師の診断を正解とした精度検証を行い、大規模出生コホートへのフェノタイピングの適用可能性を検討した。

(2) HDPの高精度な予測のための機械学習モデルの開発

フェノタイピングアルゴリズムで得られた病型分類にもとづく、HDPの予測を行う機械学習モデルの開発を行なった。入力データとしては、東北メディカル・メガバンク計画三世代コホート調査で収集されたクリーニング前の妊娠初期・中期の調査票データを利用するとともに、欠損値解析や前処理の手法について適切な手法を検討した。学習モデルにはロジスティック回帰分類器を含むいくつかの手法を適用し、それぞれの学習精度を比較した。

4. 研究成果

(1) HDPのフェノタイピングアルゴリズムの開発

本研究では国際的なHDPの診断基準に従うルールベースのフェノタイピングアルゴリズムを開発した。開発したフェノタイピングアルゴリズムを三世代コホート調査対象者へ適用することで、アルゴリズムが正しく動作することを確かめた。フェノタイピングにより得られた病型分類の結果と医師の診断を比較した結果、大部分が一致したことから、大規模出生コホートに対してフェノタイピングアルゴリズムによる病型分類が可能であることを確かめた。

(2) HDPの高精度な予測のための機械学習モデルの開発

本研究で開発した学習モデルは高い分類性能を示したことから、検討を行った前処理等の手法が適切であることが示唆された。また、大規模出生コホート調査でのHDPの早期診断支援のための機械学習の可能性が示唆された。

(3) 研究成果のまとめ

本研究では、生理学検査や尿検査、既往歴などの多くの情報源を統合的に解析し HDP の病態分類を行う高精度なフェノタイピングアルゴリズムの開発と、得られた病態分類の結果を正解とし、機械学習モデルの検討を行い、大規模出生コホート調査での HDP の早期診断支援のための機械学習構築のためのワークフローを確立した。本研究では、三世代コホート調査で収集されたクリーニング前のデータを利用し、大規模出生コホートでの HDP のフェノタイピングと早期診断支援のための機械学習の可能性の検討を行った。そのため、今後クリーニングが行われたデータでの詳細な検討を行い、三世代コホート調査での HDP 研究のための情報基盤構築を進める。

5 . 主な発表論文等

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 12 件)

水野聖土、和形麻衣子、永家聖、田宮元、栗山進一、田中博、八重樫伸生、菅原準一、荻島創一. 妊娠高血圧症候群 (HDP) 患者の病型分類のためのフェノタイピングアルゴリズムの開発と大規模前向きゲノムコホートへの適用 第 39 回日本妊娠高血圧学会学術集会, 2018.11.2~3 (大阪府)

Mizuno S, Wagata M, Nagaie S, Tamiya G, Kuriyama S, Tanaka H, Yaegashi N, Sugawara J, Ogishima S. Development of a phenotyping algorithm for Hypertensive Disorders of Pregnancy (HDP) in large-scale prospective genome cohort study. ISSHP2018, Oct 6-9, 2018 (Amsterdam)

水野聖土. 妊娠高血圧症候群の早期診断支援へ向けた機械学習の開発と知識ベースの構築. 第 1 回生命情報学研究, 2018.8.31 (東京都)

水野聖土、菅原準一、八重樫伸生. 妊娠高血圧症候群 (HDP) フェノタイピングのためのアルゴリズムの検討. 第 70 回日本産科婦人科学会, 2018.5.10~13 (宮城県)

水野聖土、永家聖、飯田溪太、笠原直子、田宮元、栗山進一、八重樫伸生、田中博、菅原準一、荻島創一. 妊娠高血圧症候群 (HDP) フェノタイピングのためのアルゴリズムの検討. 第 28 回日本疫学会, 2018.2.1~3 (福島県)

水野聖土. 妊娠高血圧症候群の早期診断支援のための機械学習とゲノム情報の利活用. 第 31 回人工知能学会, 2017.5.23~26 (愛知県)

水野聖土、田宮元、栗山進一、八重樫伸生、田中博、菅原準一、荻島創一. 妊娠高血圧症候群の早期診断支援のための機械学習とゲノム情報の利活用. NGS 現場の会第 5 回大会, 2017.5.22~24 (宮城県)

6 . 研究組織

(1)研究分担者

研究分担者氏名：なし

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号 (8 桁)：

(2)研究協力者

研究協力者氏名：なし

ローマ字氏名：

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。