

マルコフ依存混合モデルを用いた 妊娠高血圧症候群の発症リスク予測

浅井 義之^{1,2,a)} 安部 武志^{2,b)} 品川 征大^{1,c)} 前川 亮^{1,d)} 杉野 法広^{1,e)}

概要: 妊娠高血圧症候群 (HDP: Hypertensive Disorder in Pregnancy) は、妊娠中の妊婦のうち約 5% の妊婦で発症する周産期合併症の一種である。HDP は、発症転機が突然であり、不可逆的かつ急激に増悪し、母体の肝機能障害や腎機能障害、子宮内胎児発育不全など様々な障害を引き起こす。そのため、HDP の発症の早期発見と早期治療開始が求められているが、現状ではその定まった手法は確立していない。本研究では、マルコフ依存混合モデルを用い、4038 妊娠 (うち HDP278 妊娠) の妊婦健診データを用いて早期リスク予測を試みた。それぞれの妊婦検診における次回健診時の状態を推定する遷移確率から、次回以後の健診時の HDP の発症率を出力する判別器を作成した。検証データ (871 妊娠、うち HDP81 妊娠) において、妊娠中の HDP 発症予測への応用可能性が示唆された (検証 ROC-AUC 0.73)。

1. はじめに

妊娠高血圧症候群 (HDP: Hypertensive Disorder in Pregnancy) は、妊娠中の妊婦のうち約 5% の妊婦で発症する周産期合併症の一種である。HDP は、発症転機が突然であり、不可逆的かつ急激に増悪し、母体の肝機能障害や腎機能障害、子宮内胎児発育不全など様々な障害を引き起こす。重症化すると、母体痙攣である子癇や、常位胎盤早期剥離を引き起こし、妊産婦や新生児の死亡、出生児の後遺障害に大きく関与することが知られている。そのため、HDP の発症の早期発見と早期治療開始が求められている [1]。

HDP は発症転機が突然で急激に発症するという特徴がある。日本産婦人科学会診療ガイドラインによると [2]、収縮期血圧が 140mmHg 以上 (重症では 160 mmHg 以上)、あるいは拡張期血圧が 90mmHg 以上 (重症では 110mmHg 以上) になった場合、高血圧が発症したと診断される。周産期合併症の発症をできるだけ早期に捉えることを目的として、本邦では充実した頻繁な妊婦健診を実施する制度が整っており、妊娠期間中は全妊婦が密な定期的な健診を受けている。これは世界に誇れる希有な制度である。しか

し、この密な健診システムを持ってしても急激な体調変化をリアルタイムに捉えることが難しいのが現状である。妊娠初期に将来的な HDP 発症リスクを予測できる技術が必要とされている。

2. 方法

2.1 データ

疾患発症予測モデルの構築には、2009 年 10 月から 2017 年 12 月までに山口大学医学部附属病院で妊娠管理を行った 4038 例 (非 HDP 発症 3760 例、HDP 発症 278 例) を対象とした。構築した発症予測モデルの検証には、2018 年 1 月から 2020 年 3 月 31 日までに同院で妊娠管理を行った 847 例 (非 HDP 発症 766 例、HDP 発症 81 例) を対象とした。本研究は山口大学医学部附属病院臨床研究センターの承認を得て行われた。

2.2 モデル

隠れマルコフモデル (HMM) を用いた統計的機械学習を採用した。マルコフ過程で表される隠れ状態は依存する応答変数を通じて観測されるというモデルであるため、補完などに頼ることなくデータの欠損を扱え、妊婦健診のタイミングが妊婦ごとに多少の前後があっても統一して取り扱うことが可能であるという利点がある。また、Recurrent Neural Network など他の時系列データを扱う AI と比較して、HMM は学習に要する時間が短く、応答変数を除き教師ラベルを必要としない点も利点である。

本研究では、時間 (即ち妊娠週数) を共変量として状態

¹ 山口大学大学院医学系研究科

山口県宇部市南小串 1 丁目 1 - 1

² 山口大学大学院医学系研究科、山口大学医学部附属病院

AI システム医学・医療研究教育センター

山口県宇部市南小串 1 丁目 1 - 1

a) asai@yamaguchi-u.ac.jp

b) t.abe@yamaguchi-u.ac.jp

c) mshina0225@gmail.com

d) rmaekawa@yamaguchi-u.ac.jp

e) sugino@yamaguchi-u.ac.jp

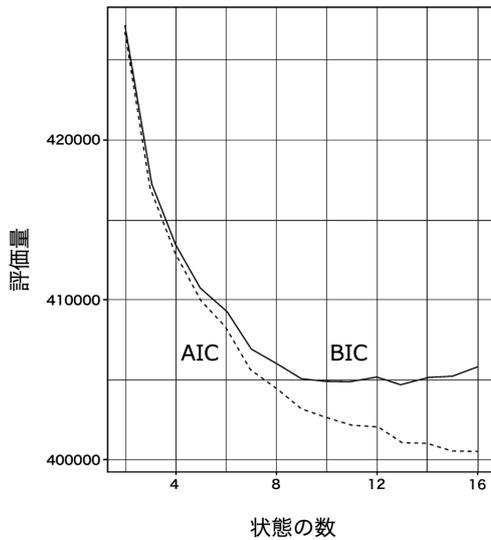


図 1 隠れ状態数決定のための情報量曲線。AIC: 赤池情報量基準、BIC: ベイズ情報量規準

遷移行列が時間依存的に変化するように（マルコフ依存混合モデル: Markov-dependent mixture model）学習を行った。これは、健常であっても妊娠中期には血圧が低下する傾向にあることを考慮するためである。収縮期血圧、拡張期血圧、尿蛋白の3つを出力変数として用いた。とり得る内部状態の数をハイパーパラメータとして扱い、10から16までの範囲で学習し、それぞれにつき情報量規準（AIC, BIC）を求め最適内部状態数を決定した。解析には、RのdepmixS4パッケージを用いた[3]。

2.3 検証

847 妊娠（うち HDP 発症 81 例）を検証データに用いて予測精度を評価した。予測に用いたスコアは、上の血圧が 140mmHg 以上または下の血圧が 90mmHg 以上となる状態についての混合確率を forward-backward アルゴリズムで求め、その妊娠全期間における最大値とした。このスコアから実際の HDP 発症の有無との比較を行い正否を決定した。

3. 結果

学習用の全妊婦健診時のデータについて、図 1 に示す情報量規準に従い、取りうる内部状態数を 14 とした。各妊娠週数における、HMM の遷移確率、即ち、ある健診の時点において、次回健診時に予測される段階が遷移確率として算出された。

予測に用いたスコアは、各妊娠週数において血圧が基準値以上に達する状態についての混合確率の最大値と定めた。このスコアから実際の HDP 発症の有無と比較した場合の予測精度の ROC 曲線を図 2 に示す。学習データについての曲線（A）と検証データについての曲線（B）はいずれも

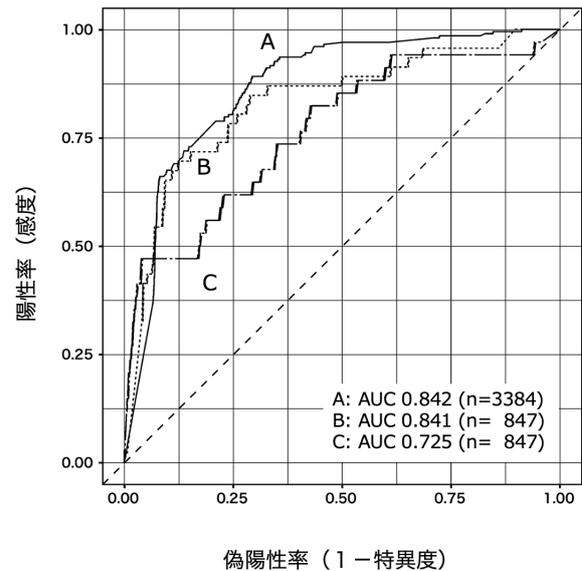


図 2 HDP 発症を判別した予測精度の ROC 曲線。A: 学習用データ（3384 人の妊婦それぞれの妊娠全期間の健診データ）の ROC 曲線。B: 検証用データ（847 人の妊婦それぞれの妊娠全期間の健診データ）の ROC 曲線。C: 検証用データ（847 人の妊婦それぞれの妊娠週数第 31 週までの健診データ）の ROC 曲線。

AUC 0.84 程度と良好な精度を示した。また、検証データを用いて、妊娠週数 31 週までの健診データのみを用いて予測した場合の曲線（C）は AUC 0.73 程度と、妊娠 31 週以前のデータからその後の HDP 発症予測についても一定の可能性が示された。

4. 結語

マルコフ依存混合モデルと妊婦健診データを用いて、妊娠早期においてその後の HDP リスク予測を行うシステム開発の可能性を示唆した。ただし、更なる精度の向上には必要である。精度向上のために、妊娠前ないし妊娠時点での妊婦の基礎的な情報を用いることで、内部状態初期分布の推定を導入することが考えられる。今後それらの技術に基づいて定量的に妊娠当初から適切なリスク層別化ができることが理想的である。良質な妊婦データを用いることにより、将来的な個別化予防医療の実現に向け、本課題は基盤となる研究である。

参考文献

- [1] Yang, Lin et al. "Predictive models of hypertensive disorders in pregnancy based on support vector machine algorithm." *Technology and health care : official journal of the European Society for Engineering and Medicine* vol. 28,S1 (2020): 181-186. doi:10.3233/THC-209018
- [2] 産婦人科診療ガイドライン 一産科編 2020. 日本産科婦人科学会, 日本産科婦人科医会, https://www.jsog.or.jp/activity/pdf/gL_sanka_2020.pdf
- [3] Visser, Ingmar et al. "depmixS4: An R Package for Hidden Markov Models." *Journal of Statistical Software* 36 (2010): 1-21. doi:10.18637/jss.v036.i07