

## 183. AI を用いた深層学習による分娩予後予測モデルの作成

森川 守

関西医科大学 医学部 産科学・婦人科学講座

Key words : 人工知能 (AI), パルトグラム, 経膈分娩, 予測モデル, 子宮口開大度

### 緒言

分娩開始前後に経膈分娩成功率ならびに経膈分娩完遂時刻（分娩所要時間）を予測できると分娩管理は現在より安心・安全に行える可能性が高い。また、最終的に経膈分娩が困難で帝王切開への変更となる確率が高いことが予測された場合、早期に変更することにより、不要な分娩誘発を中止することが可能になる。これは、安全性を向上するだけでなく、産科医・助産師ならびに妊婦の負担軽減、医療費の削減にもつながることが期待できる。さらに、妊婦ならびにその家族にその後の分娩の見通し（予測）を提示できるため、妊婦ならびにその家族に安心を与えることが期待できる。

しかし、分娩の進行は「分娩の3要素」である、娩出物（胎児、胎盤、臍帯、羊水）、娩出力（陣痛）、産道（骨盤、子宮頸管や膈）の他に、母体背景（年齢、身長や体重、既往分娩回数、分娩週数など）にも左右される。これらの複雑な多因子を医師や助産師が自己能力で全て客観的に考慮し、経膈分娩成功率ならびに経膈分娩完遂時刻（分娩所要時間）を正確に予測することは困難であり、高度な分娩予後予測モデルが必要である。その際、通常の統計学的解析のみでは困難であり、人工知能（artificial intelligence : AI）を活用し、あらゆる多様な分娩進行経過パターンに対応できる予測モデルを作成することを目的とした（図1）。

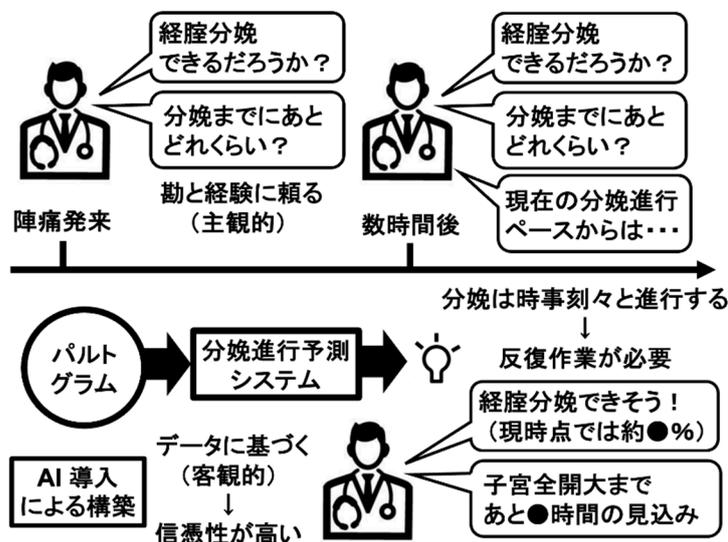


図1. 分娩進行予測の現状と解決策

現状では経膈分娩成功率ならびに経膈分娩完遂時刻（分娩所要時間）を正確に予測することは困難である。その課題を解決するためには客観的に評価する高度な分娩予後予測モデルが必要である。しかし、予測モデルを作成するためには、通常の統計学的解析のみでは困難であり、AIを活用する。あらゆる多様な分娩進行経過パターンに対応できる予測モデルを作成できると考え、本研究を計画した。

## 方法

著者の所属施設で2016年1月1日から2021年12月31日までの間に妊娠36週以降に分娩した妊婦を対象とした。1) 日本産科婦人科学会周産期登録から背景ならびに周産期予後(母体情報、新生児情報、分娩週数、分娩様式、分娩時出血量)についてのデータを収集した。2) パルトグラム(分娩経過図)(図2)ならびに胎児心拍数陣痛図から分娩経過の詳細な進行パターンについてのデータを収集した。3) 助産記録から分娩経過の概略、陣痛促進の詳細についてのデータを収集した。データは電子媒体として抽出し、データを統合・蓄積し、人工知能(AI)で解析し、深層学習 Deep learning によって分娩進行予測システムを構築した。

パルトグラム(分娩経過図)(図2)に記録された分娩経過を用いて、その後の分娩様式ならびに分娩所要時間を予測可能か評価した。パルトグラムに記録された分娩経過ならびに患者背景などの記録をもとに分娩経過を後方視的に検討し、AIを用いて深層学習させ、時系列モデルを用いた経膈分娩成功を予測するモデルを作成する。予測モデルを作成する際に、パルトグラムで記録されたCSVデータを用いて検討した。

主要評価項目は、経膈分娩予定者における経膈分娩達成(あるいは帝王切開)の予測的中率とする。その際、分娩所要時間の予測的中(誤差)も評価する。副次的評価項目は経膈分娩を達成できた要因、吸引分娩や帝王切開に至った要因とする。

関西医科大学附属病院研究倫理審査委員会の承認(No. 2021162)を得て行った。

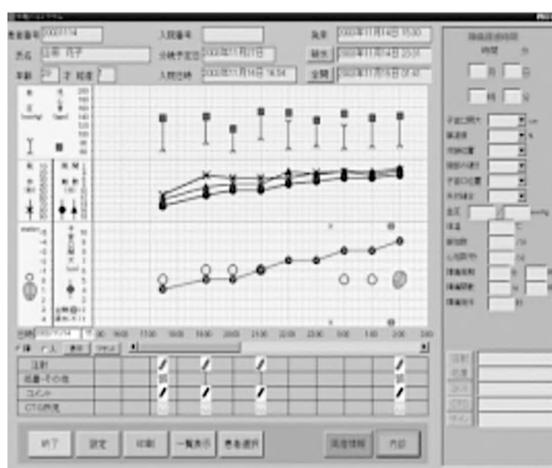


図2. パルトグラム(分娩経過図)

上部: 年齢、分娩回数、分娩予定日、入院時間、陣痛発来時間、破水時間、子宮口全開大時間

上段: 母体血圧、母体脈拍、胎児心拍数

中段: 陣痛発作時間、陣痛周期、間歇時間

下段: 子宮口開大度、児頭降下度、児頭回旋、

右部: 入力時間ごとの各入力データ、

別画面: 児娩出時間、胎盤娩出時間、子宮口展退度、児頭降下度、子宮口硬度、子宮口位置、等

これらのデータはCSVデータとして入手可能

## 結果

### 1. 研究対象の背景

関西医科大学附属病院で2016年1月1日から2021年12月31日までの間に妊娠36週以降に分娩した単胎妊婦のうち、パルトグラムのデータがなかった247例、選択的帝王切開を施行された898例、陣痛発来しな

かった 719 例を除外した 1,949 例を対象とした。なお、経膣分娩成功群が 1,729 例、不成功群が 220 例で、経膣分娩完遂率は 88.7%であった。

対象の背景を表 1 に示す。単変量解析では 2 群間に統計的に有意な多く差異を認めた。しかし、多変量解析では経膣分娩成功群の分娩時週数のみが、不成功群より有意に早かった ( $39.5 \pm 1.2$  週 vs.  $39.9 \pm 1.4$  週、 $p=0.0385$ )。なお、その差は平均で 2.3 日と大きくはなかった。

表 1. 対象の背景

	全体 (1,949 例)	経膣分娩成功 (1,729 例)	経膣分娩不成功 (220 例)	P-value
年齢 (歳)	$32.3 \pm 6.4$	$32.0 \pm 6.4$	$34.7 \pm 6.0$	<0.0001
初産 (%)	1,290 (66.2%)	1,090 (63.0%)	200 (90.9%)	<0.0001
体外受精後妊娠 (%)	365 (18.7%)	279 (16.1%)	86 (39.1%)	<0.0001
身長 (cm)	$158.9 \pm 5.3$	$158.9 \pm 5.3$	$158.5 \pm 5.7$	0.2201
妊娠前 BMI	$21.8 \pm 4.1$	$21.7 \pm 3.9$	$22.9 \pm 4.8$	<0.0001
分娩時 BMI	$25.7 \pm 4.0$	$25.5 \pm 3.8$	$27.3 \pm 4.7$	<0.0001
分娩週数 (週)	$39.6 \pm 1.2$	$39.5 \pm 1.2$	$39.9 \pm 1.4$	0.0002
分娩誘発 (%)	1,171 (60.1%)	978 (56.6%)	193 (87.7%)	<0.0001
妊娠高血圧症候群 (%)	181 (9.3%)	131 (7.6%)	50 (22.7%)	<0.0001
胎児発育不全 (%)	97 (5.0%)	87 (5.0%)	10 (4.5%)	0.8700
胎児機能不全 (%)	86 (4.4%)	32 (1.9%)	54 (24.5%)	<0.0001
無痛分娩 (%)	405 (20.8%)	361 (20.9%)	44 (20.0%)	0.8599
分娩時出血量 (mL)	$799 \pm 489$	$792 \pm 493$	$854 \pm 456$	0.0774
児出生体重 (g)	$3,070 \pm 420$	$3,054 \pm 409$	$3,190 \pm 482$	<0.0001
アプガースコア (5 分値) < 8 点 (%)	55 (2.8%)	41 (2.4%)	14 (6.4%)	0.0035
臍帯動脈血 pH	$7.32 \pm 0.08$	$7.32 \pm 0.07$	$7.27 \pm 0.10$	<0.0001

データは平均±標準偏差または比率で示した。多変量解析では経膣分娩成功群の分娩時週数のみが、不成功群より有意に早かった。

## 2. AI による予測モデルの作成過程

Bishop スコア以外、各時刻における測定値（子宮口開大、展退度、児頭位置、頸部の硬さ、子宮口位置、矢状縫合大泉門の位置、最高血圧、最低血圧、体温、脈拍数、児心拍数 1、児心拍数 2、陣痛周期、陣痛間歇、陣痛発作）を使用して、経膣分娩成功群と不成功群の二値分類を行った。

双方向 LSTM などの技術を用いた初期調査により、特定の条件下（例：開口大 8 cm 以上）で、予測精度が 80% を超えることを交差検証し確認した。過学習を防ぐためにノイズ付加等の一般的なデータ拡張手法に加え、Mixup などの画像学習で用いられるデータ拡張手法も活用した。損失関数として交差エントロピーを採用し、オプティマイザーとして Adam を用いた学習モデルの構築に成功した。学習率を学習エポックの進行に合わせスケジューリングし学習の安定性が向上した。臨床現場における予測モデルの有用性調査として、N (cm) 以上のデータをマスキングすることで、N-1 (cm) 以下の開口大について 1 cm 間隔で深層学習モデルによる解析を行った場合でも、5 cm 前後の開口大でも AUC 0.7 を超える予測が行えることを明らかにした。

## 3. AI を用いた分娩開始後に最終的に緊急帝王切開となった妊婦の予測

表 2 に分娩開始後に最終的に緊急帝王切開となった妊婦を予測する精度を示す。子宮口開大度が 4 cm から

10 cm に進行するに伴い、分娩開始後に最終的に経膈分娩に至った妊婦（成功群）を、より高い感度、特異度、陽性的中率、および Receiver Operating Characteristic (ROC) curve の area under the curve (AUC) で予測した。特に 10 cm 開大以降の分娩第 2 期で顕著であった。

表 2. 分娩開始後に最終的に緊急帝王切開となった妊婦を予測する精度

	子宮口開大度						
	4 cm	5 cm	6 cm	7 cm	8 cm	9 cm	10 cm
Cutoff 値	0.500	0.500	0.500	0.500	0.500	0.500	0.500
Average acc.	0.809	0.810	0.815	0.809	0.805	0.808	0.938
感度	0.505	0.514	0.545	0.573	0.614	0.636	0.891
特異度	0.848	0.848	0.850	0.839	0.829	0.829	0.944
PPV	0.297	0.301	0.316	0.311	0.314	0.322	0.671
NPV	0.931	0.932	0.936	0.939	0.944	0.947	0.986
AUC	0.741	0.749	0.770	0.787	0.791	0.801	0.963
95%信頼区間 (high)	0.702	0.710	0.732	0.750	0.754	0.765	0.945
95%信頼区間 (low)	0.779	0.787	0.807	0.824	0.827	0.837	0.980

Acc : accuracy、PPV : positive predictive value、NPV : negative predictive value、AUC、area under the curve、CI : confidence interval。

図 3 に、子宮口開大 N cm のタイミングで分娩後経膈分娩成功を予測するための ROC 曲線を示す。

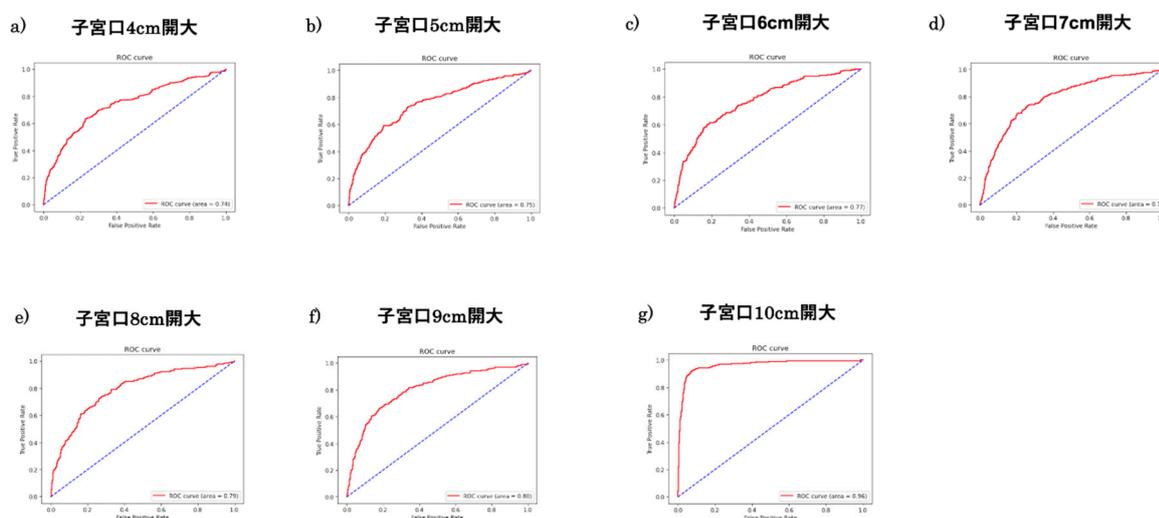


図 3. 子宮口開大度別の分娩開始後に最終的に緊急帝王切開に至った妊婦を予測するための ROC 曲線  
横軸：偽陽性率、縦軸：真陽性率、赤線：ROC 曲線。ROC : Receiver Operating Characteristic curve。

分娩誘発を行わなかった初産婦および経産婦のうち、分娩開始後に最終的に経膈分娩成功した妊婦の予測の精度も検討した。全体として、感度、特異度、PPV、NPV、および ROC curve の AUC は、子宮口開大度が 4 cm から 10 cm に進行することに伴い上昇した。しかし、結果には一貫性がなく、多少のばらつきを認めた。一方、初産婦と経産婦の間で感度、特異度、PPV、NPV、および ROC curve の AUC の違いを特定できなかった (図表の提示なし)。

## 考 察

PubMed ならびに MEDLINE を用いた検索では、AI を用いた分娩中の胎児心拍数陣痛図の判定ならびに分娩様式の予測に関する英文論文は、2018 年以降増加しているが、AI ならびに深層学習による予測が有用であるという既報論文はまだ少ない。

本研究に近い Study design の研究では、1) 前述の Guedalia J らの報告 (2020 年) [1] では、分娩室に入室時に入手可能なデータ、その後の子宮口開大度、胎児心拍数のデータを統合して作成された AI によるモデルが経膈分娩成功の予測精度を大幅に向上した。ただし、Bishop スコアは用いていない。2) Hu T らの報告 (2022 年) [2] ではオキシトシン、D'Souza R らの報告 (2023 年) [3] ではジノプロストン土オキシトシンによる分娩誘発において AI を用いて分娩誘発 (経膈分娩) 成功の予測モデルを作成し予測精度が高かった。本研究では、AI を用いて分娩誘発の有無に関わらず機能するモデルの有無にかかわらず、AI を用いて分娩誘発なしで経膈分娩が成功する予測モデルを作成した。さらに、機械学習を用いて分娩誘発を伴わない経膈分娩の成功予測モデルを作成した。しかし、症例数が少なかったために、経膈分娩の成功予測精度が低かった (結果は未表示)。なお、これらの既報研究では患者背景や分娩誘発開始時の Bishop スコアから不成功に至る危険因子を導き出し、経膈分娩成功の予測モデルを作成しているが、分娩進行中の Bishop スコアの推移は経膈分娩成功の予測には使用されていない。本研究においても、分娩進行予測モデル作成に Bishop スコアを用いたが、結果を子宮口開大度別で表示した。今後は、Bishop スコアの点数別での分娩進行予測モデルの作成を検討している。

なお、本研究での当初の目的である、分娩進行中のある時点 (内診施行時) での所見から、子宮口全開大に至るまでの所要時間の予測できるモデルの作成には至っていない。今後引き続き、検討を行う予定である。

本研究において作成した分娩進行予測モデルの精度を向上させるためには、より多くの症例数ならびに新たな追加パラメータを用いた検討が必要である。また、多施設研究による施設間のバイアスの除去が必要である。現在、引き続き、関西医科大学附属病院を研究代表施設として、同大学附属病院と同じメーカーのパートグラムを使用している大阪公立大学、奈良県立医科大学、兵庫医科大学、兵庫県立尼崎総合医療センター、市立豊中病院との多施設共同による後方視的検討が進行中である。したがって、今回報告した結果は、現時点では多施設共同研究へ向けての単施設でのパイロットスタディーという位置付けになる。

結論として、AI を用いた分娩進行に関する予測システムは、分娩開始後の経膈分娩成功の可能性を予測するのに有益であった。

今後は多施設共同研究でさらなる研究の発展を目指したい。

## 謝 辞

本研究において、AI によるデータ分析に協力いただいた株式会社イーグロース (<https://www.egrowth.co.jp/>) の今西勁峰 (代表取締役社長) 氏に謝意を表す。また、パートグラムのデータ抽出をサポートいただいた株式会社トイツー (<https://ja.toitu.co.jp/>) に謝意を表す。

## 文 献

- 1) Guedalia J, Lipschuetz M, Novoselsky-Persky M, Cohen SM, Rottenstreich A, Levin G, Yagel S, Unger R, Sompolinsky Y. Real-time data analysis using a machine learning model significantly improves prediction of successful vaginal deliveries. *Am J Obstet Gynecol.* 2020;223(3):437.e1-437.e15. doi: 10.1016/j.ajog.2020.05.025

- 2) Hu T, Du S, Li X, Yang F, Zhang S, Yi J, Xiao B, Li T, He L. Establishment of a model for predicting the outcome of induced labor in full-term pregnancy based on machine learning algorithm. *Sci Rep.* 2022;12(1):19063. doi: 10.1038/s41598-022-21954-2.
- 3) D'Souza R, Doyle O, Miller H, Pillai N, Angehrn Z, Li P, Ispas-Jouron S. Prediction of successful labor induction in persons with a low bishop score using machine learning: secondary analysis of two randomized controlled trials. *Birth.* 2023;50(1):234-243. doi: 10.1111/birt.12691.