

心理・背景因子データからの妊娠期うつ病発症リスク診断

富田 康之¹・加藤 竜司¹・野坂 幸世¹・武藤 裕紀¹・富田真紀子²・濱田 祥子²
涌井 陽子²・金井 篤子²・村瀬 聡美²・臼井比奈子³・石川 直子³・廣瀬 允美³
水野 妙子³・後藤 節子³・尾崎 紀夫³・本多 裕之^{1,4*}

(2008年7月25日受付 2008年10月6日受理)

Risk Prediction of Depression Crisis during Pregnancy Using Supervised Learning Model

Yasuyuki Tomita¹, Ryuji Kato¹, Yukiyo Nosaka¹, Hironori Mutoh¹, Makiko Tomida², Shoko Hamada², Yoko Wakui², Atsuko Kanai², Satomi Murase², Hinako Usui³, Naoko Ishikawa³, Masami Hirose³, Taeko Mizuno³, Setsuko Gotoh³, Norio Ozaki³, and Hiroyuki Honda^{1, 4*}
(School of Engineering¹, School of Education², Nagoya University, Furocho, Chikusa-ku, Nagoya 464-8603; School of Medicine, Nagoya University, Tsurumai-cho, Showa-ku, Nagoya 466-8550³; MEXT Innovative Research Center for Preventative Medical Engineering, Furocho, Chikusa-ku, Nagoya 464-8601⁴) Seibutsu-kogaku **86**: 519-523, 2008.

Depression crisis during pregnancy is a type of mental diseases which 10 to 15 percentage of pregnant women suffers. Since depression is associated with multifactorial causes including genetic factor and acquired social conditions, it has been a challenge to analyze the incidental factors that increases the risk of depression crisis. Among other depressions, depression during pregnancy has an advantage of clear determination of causable period, therefore considered to be model case of depression analysis. In this report, the combinational analysis of risk factors using supervised learning algorithm was applied to the questionnaire based information of patients with depression risks. By the appropriate combination with stratification of samples and prediction model, we succeeded in understanding depression risk with combinational risk factors.

[**Key words:** risk prediction, depression, pregnancy, supervised learning, multifactorial]

うつ病とは、発症に伴う社会活動の低減だけでなく、周囲の人々に影響をもたらす重大な現代病である。その発症は、環境的な要因や遺伝的な要因が密接かつ複雑に関係しており、特定が非常に難しい。特にうつ病は、発症時期が特定できず、脳組織の生検も不可能であるため、病態解明がより困難であり、根本的な治療が難しい。

妊娠期うつ病は、10～15%程度の妊婦が妊娠をきっかけに、「気分が沈む、興味が失せる、くよくよ悩む」などの、軽い憂うつ感を感じるといった病状を示す症状である。その発症には多様な因子が関連しており、女性の社会進出に伴う葛藤や、心理社会的要因が大きく関わっている。他のうつ病に比べると発症時期を規定しやすい

妊娠期うつ病は現在、うつ病発症要因を解明する解析モデルとして非常に重要だと考えられている²⁻⁴⁾。また妊娠期うつ病の理解によるメンタルヘルスクアは、多くの女性をサポートするため非常に重要である。本研究では、妊娠期うつ病の関連因子解析を行い、妊娠期うつ病発症に重要な因子を特定するための有効な手法を提案することで、メンタルヘルスクアのためのリスク診断手法の確立に貢献することを目標とした。

本研究では、うつ病関連因子の解析方法として、従来の統計的アンケート解析²⁾やパラメトリックな解析手法ではなく、ノンパラメトリック解析法である学習アルゴリズム (Fuzzy Neural Network: FNN) の適応の可能性

*連絡先 ¹名古屋大学工学研究科 (〒464-8603 名古屋市千種区不老町)
TEL. 052-789-3215 FAX. 052-789-3214 E-mail: honda@nubio.nagoya-u.ac.jp

²名古屋大学大学院教育学研究科, ³名古屋大学大学院医学系研究科, ⁴名古屋大学予防早期医療創成センター

を検証した。FNNは、これまで多くの生物データ解析において重回帰分析やロジスティック回帰のような解析よりも高い予測精度を示しており⁵⁾、説明変数の組合せをルール表として記述することができる利点を持つ。このため、臨床現場においてうつ病診断を現実的に行う医師を支援する観点から、パラメトリックな解析と共に比較検証すべき一つの有効なモデル化手法として、その有用性を提案できるのではないかと考えた。

このため具体的には、妊娠初期、中期の女性を対象として取得された社会心理的要因（人格傾向、母親自身の過去の養育体験、ソーシャルサポート状況、ワークファミリーコンフリクト）と個人背景要因（仕事の有無、年齢、子供の数、家族構成）などのアンケート情報を学習したFNN予測モデルを構築し、アンケートデータのような曖昧な情報量をもつ項目の組合せ効果や層別化について検討した。

実験方法

解析データ 解析データは、名古屋大学教育発達科学研究科において、2004年7月から名古屋市内の病院で実施された「妊産期うつ病に関するアンケート調査」により収集されたデータを用いた。本研究データの解析にあたっては、名古屋大学医学部倫理委員会において承認済された内容（承認番号：513）に基づき、全対象者から文書による同意を得たものである。サンプル数（アンケート回答者）は193人（妊娠初期、中期を含む）。アンケートは、心理社会的な要因や個人背景要因などを明らかにするための127の質問から構成されている（一例をTable 1中に示す。詳細は参考文献2を参照されたい）。127の質問への回答は、因子分析の結果を受けて1因子性を持つ項目をまとめて主要17説明変数因子とし、各因子を構成する項目群の合計点を項目数で割った値をスコアとした（Table 1）。各質問は、ランダム化や逆転項目を設けるなどすることで、人為的な得点とならないように設計されている。各サンプルのうつ病の評価には、エジンバラ産後うつ病自己評価表（Edinburgh Postnatal Depression Scale; EPDS）⁶⁾による得点指標を用いた。EPDS指標値はうつ病の程度を示す国際的指標で、値が高いほうがうつ病の傾向が高く、一般には9以上でうつ病になるリスクが高いとされる。

解析方法 本解析においては、説明変数（入力値）には、スコア化された17項目のアンケート設問回答内容、従属変数（出力値）にはEPDSの評価指標の値を用いた。

解析には、ファジィ推論と人工ニューラルネットワーク（Artificial Neural Network; ANN）とを組み合わせた

Table 1. Contents of 17 questionnaire items for diagnosis of EPDS.

Questionnaire items	Abbreviation
1 Age	ag
2 Work (1: working, 0: not working)	wk
3 Relationship to the child (1: first child, 0: others)	rs
4 Work involvement	wki
5 Family involvement	fmi
6 The scale of egalitarian sex role attitudes	sxri
The expected anxiety of work-family conflict (WFC)	
7 Selective conflict	sc
8 Family-work conflict	fwc
9 Time conflict	tmc
10 Work-family conflict	wfc
11 Harm avoidance	ha
12 The numbers of social support	nm
13 Satisfaction level of social support	sss
14 Care (father)	crf
15 Care (mother)	crm
16 Overprotection (father)	opf
17 Overprotection (mother)	opm

Example of questionnaire

	ち ま が う た く	い え ば ち が う か と	ど ち ら か と も	い え ば そ の 通 り	ど ち ら か と も	ま つ た く の 通 り
1 仕事は私にとって非常に大事な一部分である	1	2	3	4	5	
2 私は自分の仕事に大変熱中している	1	2	3	4	5	
3 私は自分の仕事に深く傾倒している	1	2	3	4	5	
4 私にとって、生活上の他の事柄に比べれば、仕事はそんなに重要ではない	1	2	3	4	5	
5 仕事は私の人生の目標を達成する上で、中心的な位置を占めている	1	2	3	4	5	
6 たとえ自分がたくさんのお金を持ち、生活の維持を心配する必要がなくなったとしても、私は働く	1	2	3	4	5	

FNN^{5,7)}を用いた。FNNは、単なる予測モデルとしてではなく、学習（モデル内の各パラメータの自動最適化）の過程を経て得られたモデル構造から説明変数と従属変数の関係を最もよく記述できるルールを発見できる。

解析は、(1) 193人のアンケートスコアを用いたEPDSの値（連続値）を予測する連続値予測モデルの構築と、(2) 193人中57人のアンケートスコアを用いた、EPDS

Table 2. Result of the selected input variables among top EPDS prediction models.

	Ranking	Input variables														Number of inputs				
		ag	wk	rs	wki	fmi	sxri	sc	fvc	tmc	wfc	ha	nm	sss	crf		crm	opf	opm	
Continuous prediction model	1							○				○		○						3
	2								○			○		○						3
	3											○		○					○	3
	4				○							○		○						3
	5											○		○					○	3
	6	○										○		○						3
	7				○				○			○		○						3
	8											○		○		○				3
	9				○							○		○		○				3
	10				○							○		○					○	3
Binary prediction model	1										○				○				○	3
	2										○				○					3
	3										○			○						3
	4				○						○			○					○	3
	5					○					○			○						3
	6					○					○			○					○	3
	7							○			○			○			○			3
	8							○			○			○					○	3
	9								○		○			○						3
	10					○					○			○			○			3

Abbreviations of input variables are indicated in Table 1.

値0の群 (29人)・EPDS値9以上の群 (28人) の2群判別モデルの構築を行った。特に後者では、EPDSの正常範囲が広く、うつ病の症状自体にも明確な境界はないため、「うつ症状のないサンプル」としてEPDS値が0のサンプルを用いた。

各解析では3-foldクロスバリデーションを用いた。各モデル構築における学習過程においては、モデルの評価基準を設定し、これを最適化するようにモデルの更新が行われる。(1)連続値予測モデル構築では、実際のEPDSの値(教師値)とテストデータにおいて推定されたEPDSの値(推定値)の平均二乗平方根誤差(Root mean square error; RMSE)を算出し評価値として予測モデルを評価した。(2)2群判別モデルの構築では、テストデータにおいて正答率(Accuracy)、感度(Sensitivity)、特異度(Specificity)の3つを予測モデルの評価値として用いた。教師値をEPDS値0の群を-1、EPDS値9以上の群を+1に設定し、モデルから得られる推定値が0より大きい(case)、小さい(control)かの判定を行い、正答率((True positiveの数+True negativeの数)/(Caseの数+Controlの数))、感度(True positiveの数/Caseの数)、特異度(True negativeの数/Controlの数)を算出した。

結 果

EPDS連続値予測モデルによる解析 アンケート項目(1~3項目)を用いた全組合せ(833通り)に対し、EPDS値を連続値として予測するモデルをすべて構築し、予測精度の高い順番にランキングを行い、選択された変数の組合せを検証した(Table 2)。結果、最も予測

精度の良いモデルは(1)職業か家族かを選択しなくてはいけない葛藤による不安(Selective conflict; sc)、(2)損害回避傾向(Harm avoidance; ha)、(3)社会的支援に関する満足度(Satisfaction level of social support; sss)の3つの組合せを用いたものであり、同時にhaは上位モデルに共通して選択されていたため、特に重要であることが示唆されていた。

次に、最もRMSEの小さなモデルの予測結果をFig. 1に示した。横軸には実際に各患者に対して算出されていたEPDS値、縦軸に予想モデルより算出された推定EPDS値をプロットした。結果、予測誤差((教師値-推定値)の絶対値)が2以下の場合、または3以下の場合でも、正答率はそれぞれ47%(193人中91人正解)、65%(193

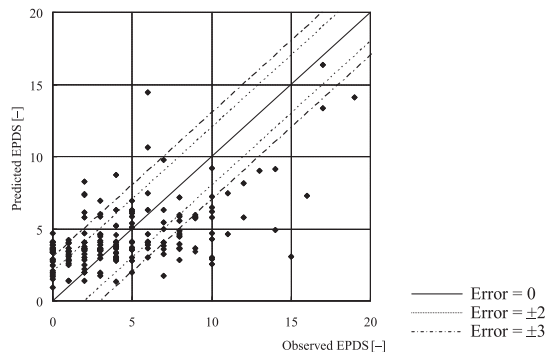


Fig. 1. Relationship between observed EPDS and predicted EPDS by the FNN model which RMSE in test data was the lowest in all models within 3 input variables. Circled samples whose EPDS values were more than 16 were predicted more accurately compared to other samples with serious depression. The 3 input variables were selective conflict (sc), harm avoidance (ha), and satisfaction level of social support (sss).

人中126人正解)であった。すなわち、17項目中3項目の組合わせの妊産期うつ病への関連性は相対的に明らかであったが、各患者のEPDSの連続値としての予測精度は低かった。比較対象として重回帰分析を行ったが、予測精度は最良モデル(4入力: ha, sss, prm, sc)であっても42.5% (誤差2以下), 61.7% (誤差3以下)とFNNを下回っていた。

EPDSによる2群分類モデルによる解析 前検討において、全サンプル・最大3説明変数を用いた予測モデルでは予測精度があまり向上しなかったことから、妊産期うつ病リスクの診断モデルとしての実用性を向上させるため、2群判別モデルの構築を検討した。特に「うつ発症リスクの高いサンプル」をいかに高精度に予測できるかに重点をしぼり、全サンプルを実際のEPDSの値が9以上(うつ程度が高い)サンプルと実際のEPDS値が0(うつ症状のない)サンプルの2群に層別化し判別分類を行った。

新たに、アンケート項目(1~3項目)を用いた全組合せ(833通り)に対し2群判別モデルを構築し、評価データにおける予測精度の高さでランキングを行い、説明変数の組合せを検証した(Table 2)。結果、予測精度が最も高いモデルは、損害回避傾向(Harm avoidance; ha)、父親の心遣い度合(Care (father); crf)、母親の過保護度合(Overprotection (mother); opm)の3つの組合せを用いたモデルであった。ここでも連続値予測モデルと同様にhaが説明変数として多数選ばれ、重要な因子であることが示唆されていた。モデルが変更された場合においても説明変数として何度も選ばれる項目は、妊産期うつ病発症リスクの非常に高いグループと無いグループとを識別するのに重要な項目である可能性が高い。Table 3には、予測精度の最も高いモデルを例に予測結果の詳細を示した。結果、連続値予測モデルとは異なり、実際のEPDSの値が9以上のサンプルにおいても高精度の予測が実現されていた。比較対象としてロジスティック回帰分析を行ったが、予測精度は最良モデル(5入力: sc, wfc, ha, sss, prm)であっても86.2%とFNNを下回っていた。

		Care (father)			
		low		high	
		Overprotection (mother)			
		low	high	low	high
Harm Avoidance	low	1 / 0	0 / 2	0 / 6	0 / 2
	high	3 / 2	8 / 7	6 / 9	10 / 1

Fig. 2. Fuzzy rule interpreted from the best binary prediction model. Each cell indicates the combinatorial condition of 3 input parameters, such as care (father) (crf), overprotection (mother) (opm), and harm avoidance (ha). Numbers in each cells are; (right) number of samples (EPDS = 9), and (left) number of samples (EPDS = 0). The grey colored cell indicates the parameter combination which can explain the maximum number of patients whose EPDS ≥ 9.

この2群判別予測モデルから得られた3説明変数の組合せルールをFig. 2に示した。灰色で示したセルは「crfの値が大きく、opmの値が大きく、haの値が大きい」という組合せルールを示し、うつ発症リスクの高いサンプルが偏って所属していることがわかった。

すなわち、サンプルを層別化し判別予測を行うことによって高精度の妊産期うつ病発症リスクの有無の予測が可能となり、学習モデルの構造を読み解くことによって、うつ発症リスクの高いサンプルが持つ項目の組合せに傾向を見つけることができた。

考 察

学習モデルを用いた妊産期うつ病発症関連因子の選択結果からは、連続値予測モデルと2群判別予測モデル共に、上位10モデルのほとんどでha(Harm avoidance: 損害回避傾向項目)が説明変数として選択されていた(Table 2)。これは、haに関わるアンケート設問が、妊産期うつ病発症のリスクを予測するのに非常に重要なことを強く示唆するものである。これは、既存の研究における報告と合致するものであった⁸⁾。本研究の解析対

Table 3. Result of the best model in binary prediction analysis.

Samples	The number of correctly predicted samples	The number of samples	Predictability	
EPDS ≥ 9	24	28	Sensitivity	85.7%
EPDS = 0	28	29	Specificity	96.6%
All	52	57	Accuracy	91.2%

Indicated accuracy, sensitivity, and specificity are the results of 3-fold cross-validation. The 3 input variables were harm avoidance (ha), care (father) (crf), and overprotection (mother) (opm).

象である193人では既往の研究と比較して得られたデータが極端に偏っている現象は確認されなかったが、今後のさらなるデータの蓄積によって本解析で選ばれた項目の信頼性が証明されると考えられる。

うつ病の最近の研究によれば、病前性格により、うつ病にhigh riskなものを同定することが可能となるという報告がある⁹⁾。人格に関する議論も盛んで、より遺伝性の高い人格要素も注目されている。これには「新奇性追求 (NS)」、「損害回避 (HA)」、「報酬依存 (RD)」、「固執 (P)」の4つがあり、それぞれ、NSが中枢神経内のdopamine, HAがserotonin, RDがnorepinephrineの神経伝達物質の分泌と代謝に依存しているものと想定されている¹⁰⁾。これらの知識はうつ病の予防や早期介入に重要と考えられているため、今後は上記の各伝達物質の血中濃度も組み合わせた解析、あるいは遺伝子多型を組み合わせた解析も重要であろうと思われる。

本研究で提案するノンパラメトリックなモデルのアンケートへの適応は、臨床上でより正確なリスク診断をする際、他の予測モデルと共に多面的に症状を診断しようとする際、その簡便さと理解しやすさから大いに現場の医師をサポートできると考えられる。

要 約

本研究では、妊産期うつ病の発症リスク診断法の一つとして、得られたアンケート結果をノンパラメトリックな学習モデルによって解析し、疾患関連因子の組合せの探索と診断システムとしての予測モデルの有効性を検討した。結果、うつ病のような多因子の複雑な発症因子解

析には、ノンパラメトリックな学習モデルによる解析は迅速かつ簡便であり、データの層別化によって有効な解析モデルを診断モデルの一つとして得られることを強く示すものであった。

本研究は、名古屋大学高等研究院研究プロジェクトおよび平成16年度赤崎記念研究奨励事業、平成17～19年度文部科学省科学研究補助金基盤研究 (B) 課題番号17390183、平成17～19年度文部科学省科学研究補助金基盤研究 (B) 課題番号17310151により遂行された。

文 献

- 1) Dennis, C. L., Ross, L. E., and Grigoriadis, S.: *Cochrane Database Syst. Rev.*, **3**, CD006309 (2007).
- 2) Kanai, A. and Tomida, M.: *The 21st Japanese Association of Industrial/Organizational Psychology, Saitama, Japan* (in Japanese) (2005).
- 3) Taylor, A., Atkins, R., Kumar, R., Adams, D., and Glover, V.: *Arch. Womens Ment. Health*, **8**, 45–51 (2005).
- 4) Reck, C., Klier, C. M., Pabst, K., Stehle, E., Steffenelli, U., Struben, K., and Backenstrass, M.: *Arch. Womens Ment. Health*, **9**, 265–271 (2006).
- 5) Ando, T., Suguro, M., Kobayashi, T., Seto, M., and Honda, H.: *Cancer Sci.*, **94**, 906–913 (2003).
- 6) Cox, J. L., Holden, J. M., and Sagovsky, R.: *Br. J. Psychiatry*, **150**, 782–786 (1987).
- 7) Horikawa, S., Furuhashi, T., Uchikawa, Y., and Tagawa, T.: *Proc. Int. Fuzzy Eng. Symp. '91*, p. 562–573 (1991).
- 8) Kitamura, T., Shima, S., Sugawara, M., and Toda, M. A.: *Psychol. Med.*, **23**, 967–975 (1993).
- 9) 佐藤哲哉, 上原 徹: *精神科診断学*, **6**, 399–428 (1995).
- 10) Cloninger, C. R., Svrakic, D. M., and Przybeck, T. R.: *Arch. Gen. Psychiatry*, **50**, 975–990 (1993).