

Original Article

人工ニューラルネットワークを用いた脳卒中患者の更衣自立予測モデルにおけるサンプル数と精度の関係

藤田貴昭,¹ 大橋拓朗,² 山根和広,² 山本優一,²
 曾根稔雅,¹ 大平葉子,² 大槻剛智,² 五百川和明³

¹東北福祉大学健康科学部リハビリテーション学科

²北福島医療センターリハビリテーション科

³福島県立医科大学新医療系学部設置準備室

要旨

Fujita T, Ohashi T, Yamane K, Yamamoto Y, Sone T, Ohira Y, Otsuki K, Iokawa K. Relationship between the number of samples and the accuracy of the prediction model for dressing independence using artificial neural networks in stroke patients. Jpn J Compr Rehabil Sci 2020; 11: 28-34.

【目的】人工ニューラルネットワークを用いて脳卒中患者の更衣の自立可否の予測モデルを作成する際に、有用なモデルが作成できるサンプル数の下限の目安を調べる。

【方法】脳卒中患者 121 名から無作為抽出を繰り返し 120 名, 100 名, 80 名, 60 名, 40 名の 5 つのデータセットを作成した。それぞれのデータセットで回復期リハ病棟入院時の変数から入院 1 か月後の更衣自立可否を予測するモデルを人工ニューラルネットワークとロジスティック回帰で作成し、両モデルの精度を比較した。

【結果】120 名, 100 名, 80 名のデータセットでは、人工ニューラルネットワークモデルの精度がロジスティック回帰モデルと比較して有意に高かった。一方、60 名と 40 名のデータセットでは両モデルの精度に差は認められなかった。

【結論】人工ニューラルネットワークで有用な更衣自立予測モデルが作成できる下限サンプル数は 80 程度であることが示唆された。

において対象者の日常生活活動能力 (ADL) の予後予測は、介入方針や目標の設定、必要となる人的・物理的環境の準備、および家族指導などにおいて重要となる。ADL 予後予測法については種々の報告があるが、結果の一般化に限界を有することが多く [2, 3], 予測の精度を高めるためには各施設で独自の予測モデルを作成することが重要であることが指摘されている [4]。

一方で、各単一施設で予測モデルを作成する場合には使用可能なデータ数が少数になりやすいことが想定される。近年、Fujita ら [5] は各単一施設のサンプルでも精度の高い ADL 予測モデルを作成できる手法を探索し、人工ニューラルネットワークが有用であったことを報告している。同研究では 83 例のサンプルを用いて更衣自立に関する予測モデルを作成した場合、ロジスティック回帰や決定木と比較して人工ニューラルネットワークのモデルの精度が高かったことを報告しているが、その一方で人工ニューラルネットワークの精度の優位性が保たれるサンプル数の下限は明らかにされていない。そこで本研究では、人工ニューラルネットワークで有用な更衣自立の予測モデルを作成できるサンプル数の下限の目安を調べるため、人工ニューラルネットワークによる予測モデルの精度と学習サンプル数の関係を検証した。

対象と方法

本研究は後方視的観察研究であり、対象者の診療録から種々の情報を収集して分析した。対象は A 病院回復期リハ病棟に入院した脳卒中患者で以下の取り込み基準を満たした 121 名である。取り込み基準は (1) 初発の脳出血または脳梗塞と診断された者、(2) 一側テント上の病変であった者、(3) 入院評価時に更衣が非自立であった者 (入院時 FIM[®] instrument (version 3) 更衣上半身・下半身 [6, 7] の一方または両方が 5 点以下)、(4) 後述の分析データに欠損値がなかった者とした。本研究は北福島医療センターと東北福祉大学の倫理審査委員会の審査を受け、承認を得た後に実施した (No. 72, RS180601)。

本研究では入院時の変数から入院 1 か月後の更衣自立可否を予測するモデルの作成を行った。入院 1

キーワード：脳卒中、予後予測、日常生活活動

はじめに

脳卒中は長期的な能力低下を引き起こす主要な原因疾患の一つである [1]。リハビリテーション (リハ)

著者連絡先：藤田貴昭
 東北福祉大学健康科学部リハビリテーション学科
 〒989-3201 宮城県仙台市青葉区国見ヶ丘 6-149-1
 E-mail : t-fujita@tfu-mail.tfu.ac.jp
 2020 年 1 月 7 日受理

利益相反：本研究に関連する利益相反は一切ありません。

か月後の更衣自立可否の判断は、入院1か月時のFIM[®] instrument 更衣項目の得点を指標とした。なおFIM[®] instrument 更衣項目には上半身と下半身があるが、本研究では両者のうち低いほうの得点を採用し、6点以上を自立、5点以下を非自立とした。予測モデルの独立変数として用いる入院時変数には、これまでに更衣自立度と関連することが報告されている年齢 [8]、練習前の更衣自立度 [9]、体幹機能 [8, 10]、視空間認知 [8]、バランス [11, 12] の情報を収集した。入院時の更衣自立度の指標には入院時のFIM[®] instrument 更衣項目の得点、体幹機能と視空間認知の指標には Stroke Impairment Assessment Set (SIAS) [13]、バランスの指標には Berg balance scale (BBS) [14] を用いた。またこれらの変数に加えて、本研究ではADLの予後と関連することが報告されている麻痺側および非麻痺側の上肢機能、認知機能、脳卒中発症後期間も変数に加えた。上肢機能の指標には簡易上肢機能検査 (STEF) [15]、認知機能の指標には改訂長谷川式知能評価スケール (HDS-R) [16] を用いた。

本研究では人工ニューラルネットワークを用いた更衣自立予測モデルの精度とサンプル数の関係を調べるため、対象者121名から無作為抽出を繰り返して120名、100名、80名、60名、40名の5つのデータセットを作成し、それぞれのデータセットで予測モデルを作成してその精度を検証した。人工ニューラルネットワークで作成されたモデルの精度については、一般的に広く用いられている手法であるロジスティック回帰で作成されたモデルの精度と比較することで検討した。人工ニューラルネットワークおよびロジスティック回帰によるモデル作成の手順としては、まずサンプル数の異なる5つのデータセットのそれぞれで、対象者を入院1か月時のFIM[®] instrument 更衣得点に基づき、自立群 (FIM[®] instrument ≥ 6 点) と非自立群 (FIM[®] instrument ≤ 5 点) に分類し、モデルに投入する変数の選択のため、各入院時変数について群間比較を行った。比較には student-*t* 検定、 χ^2 検定、Mann-Whitney *U* 検定を用いた。次に群間比較で有意であった入院時変数を独立変数、入院1か月時更衣の自立・非自立を従属変数として、ステップワイズロジスティック回帰分析 (尤度比変数増加法) を行い、ロジスティック回帰モデルを作成した。一方、人工ニューラルネットワークモデルは本研究では中間層1層の階層型多層パーセプトロンとし、独立変数はロジスティック回帰で選択された変数として、ロジスティック回帰と人工ニューラルネットワークのモデルが同じ独立変数で作成されるように設定した。また本研究では人工ニューラルネットワークにおける学習サンプルと過学習を防ぐための検定サンプルの割合を9:1とした。また人工ニューラルネットワークではサンプル数を考慮して、5段階を超える変数 (つまりSIAS 体幹、視空間認知以外の項目) が独立変数に使用される場合には四分位を基準に4段階に変換し投入した。なお人工ニューラルネットワークでは、入力層から中間層への重みの初期値がランダムであり、学習結果はこの初期値に依存するため、本研究では初期値の再設定を10回反復し、最も予測精度の高かったモデルを採用した。

人工ニューラルネットワークとロジスティック回帰

で作成されたモデルの精度は層化10分割交差検証により比較した。10分割交差検証は全対象者を無作為に10分割し、9つのグループを使用してモデルを作成し、残りの1つのグループを検証データとして精度を検証することを10回繰り返す方法である。この際、本研究ではロジスティック回帰ではモデルで用いられる独立変数が変化しないように強制投入法を用いた。10分割交差検証では人工ニューラルネットワークとロジスティック回帰それぞれのモデルにおける正分類率、感度、特異度、陽性的中率、陰性的中率を算出してウィルコクソンの順位和検定を用いて比較した。すべての検定の有意水準は5%未満とし、統計ソフトには SPSS Statistics version. 25 を用いた。

結果

各データセットの対象者の属性および心身機能、更衣自立度を表1に示す。入院1か月後の更衣自立群と非自立群で入院時変数を比較した結果、100名のデータセットでの年齢と麻痺側STEFを除き、すべての変数で有意差が認められた (表2)。それらの変数を独立変数としてロジスティック回帰分析を実施した結果、120名のデータセットではBBSと非麻痺側STEFとSIAS垂直性、100名のデータセットではBBSと非麻痺側STEFと発症後日数、80名のデータセットではBBSとHDS-R、60名のデータセットではBBSと非麻痺側STEFと年齢、40名のデータセットではBBSと非麻痺側STEFが、モデルの独立変数として選択された。

作成された人工ニューラルネットワークとロジスティック回帰のモデルの精度を比較した結果、120名、100名、80名のデータセットを用いた分析では、人工ニューラルネットワークのモデルが正分類率、感度、特異度、陽性的中率、陰性的中率のすべてでロジスティック回帰のモデルを上回り、120名の分析では正分類率と陰性的中率と感度、100名の分析では陽性的中率、80名の分析では正分類率で有意差が認められた (表3)。一方、60名、40名の分析ではすべての項目において人工ニューラルネットワークとロジスティック回帰のモデルの間に有意差は認められなかった。

考察

医学・医療分野において人工ニューラルネットワークを用いて予測モデルを作成し、その精度を従来の一般的な手法であるロジスティック回帰モデルと比較した報告は多数存在する。それらの報告を俯瞰すると、ロジスティック回帰モデルと人工ニューラルネットワークの間で予測性能の差はないとする報告 [17-22] が一定数存在するものの、人工ニューラルネットワークのほうが優れていると結論付けた報告 [23-31] が多い。また近年、外傷患者の転帰に関するシステマティックレビューとメタアナリシスを行った研究 [32] においても、人工ニューラルネットワークのモデルはロジスティック回帰よりも優れた性能を有することが報告されている。また著者らは83例という小サンプルであっても、脳卒中の更衣の自立を予測に関して、人工ニューラルネットワークはロジスティック

表 1. 各データセットの属性, 心身機能および更衣自立度

Variables	Dataset 1 (<i>n</i> =120)	Dataset 2 (<i>n</i> =100)	Dataset 3 (<i>n</i> =80)	Dataset 4 (<i>n</i> =60)	Dataset 5 (<i>n</i> =40)
Age, years, mean (SD)	75.2 (12.4)	75.4 (12.6)	75.6 (13.0)	74.8 (13.0)	74.8 (14.3)
Gender, men, <i>n</i> (%)	62 (51.7)	57 (57.0)	43 (53.8)	32 (53.3)	22 (55.0)
Affected side, right, <i>n</i> (%)	56 (46.7)	51 (51.0)	37 (46.3)	25 (41.7)	20 (50.0)
Post-stroke time at admission, days, mean (SD)	30.2 (11.7)	29.3 (11.2)	29.4 (11.1)	31.4 (12.0)	28.4 (11.9)
SIAS verticality at admission, points, median (IQR)	3 (2-3)	3 (2-3)	3 (2-3)	3 (2-3)	3 (1-3)
SIAS abdominal muscle strength at admission, points, median (IQR)	2 (1-2)	2 (1-2)	2 (0.5-2)	2 (1-2)	2 (1-2)
SIAS visuospatial perception at admission, points, median (IQR)	3 (3-3)	3 (3-3)	3 (3-3)	3 (3-3)	3 (3-3)
Berg balance scale at admission, points, mean (SD)	20.5 (16.8)	21.3 (16.6)	20.3 (17.1)	20.5 (16.6)	21.6 (16.3)
STEF affected side at admission, points, mean (SD)	75.1 (21.0)	76.3 (20.6)	75.2 (22.2)	75.7 (20.6)	73.1 (25.4)
STEF unaffected side at admission, points, mean (SD)	25.8 (32.8)	27.5 (33.5)	24.5 (32.9)	27.3 (32.8)	31.7 (33.9)
HDS-R at admission, points, mean (SD)	19.7 (7.9)	19.5 (7.9)	19.5 (7.9)	20.9 (7.2)	19.6 (7.6)
Dressing performance at admission					
FIM® upper body dressing, points, median (IQR)	3.0 (1.0-4.0)	3.0 (1.0-4.5)	3.0 (1.0-4.0)	3.0 (1.0-4.0)	3.0 (1.0-4.5)
FIM® lower body dressing, points, median (IQR)	2.0 (1.0-4.0)	2.0 (1.0-4.0)	2.0 (1.0-4.0)	2.0 (1.0-4.0)	2.0 (1.0-5.0)
The lower score on FIM® for dressing the upper and lower body, points, median (IQR)	2.0 (1.0-4.0)	2.0 (1.0-4.0)	2.0 (1.0-4.0)	2.0 (1.0-4.0)	2.0 (1.0-4.0)
Dressing performance at 1-month after admission					
FIM® upper body dressing, points, median (IQR)	5.0 (2.0-6.0)	5.0 (2.5-6.0)	5.0 (2.0-6.0)	4.5 (2.5-6.0)	5.0 (2.0-6.0)
FIM® lower body dressing, points, median (IQR)	4.0 (2.0-6.0)	5.0 (2.0-6.0)	4.0 (2.0-6.0)	4.5 (2.5-6.0)	4.5 (2.0-6.0)
The lower score on FIM® for dressing the upper and lower body, points, median (IQR)	4.0 (2.0-6.0)	5.0 (2.0-6.0)	4.0 (2.0-6.0)	4.0 (2.0-6.0)	4.0 (2.0-6.0)
Independence of upper body dressing, <i>n</i> (%)	39 (32.5)	34 (34.0)	26 (32.5)	20 (33.3)	14 (35.0)
Independence of lower body dressing, <i>n</i> (%)	37 (30.8)	32 (32.0)	24 (30.0)	19 (31.7)	12 (30.0)
Independence of both upper and lower body dressing, <i>n</i> (%)	37 (30.8)	32 (32.0)	24 (30.0)	19 (31.7)	12 (30.0)

Abbreviations: SIAS, stroke impairment assessment set; HDS-R, revised Hasegawa's dementia scale; STEF, simple test for evaluating hand function.

ク回帰より予測精度の高いモデルを作成できたことを報告した [5]。一方で、人工ニューラルネットワークとロジスティック回帰の精度の差に関して、サンプル数を変化させて検証した研究はこれまでになく、また 100 未満の小サンプルを用いた報告もほとんど見当たらない。そのため、人工ニューラルネットワークで有用なモデルを作成できるサンプル数の下限はこれまで不明であった。

本研究の結果、人工ニューラルネットワークで有用

な更衣自立予測モデルが作成できる下限サンプル数は 80 程度であり、サンプル数が 60 以下になるとその優位性は失われる可能性が示唆された。80 名程度のサンプルにおいて、ロジスティック回帰に比べて人工ニューラルネットワークのほうが精度の高いモデルが作成できたことは先行研究 [5] の結果と一致する。今回 100 名のデータセットでは、120 名と 80 名のデータセットと異なり、人工ニューラルネットワークとロジスティック回帰で正分類率に有意な差は認めら

表 2. 入院 1 か月後の更衣自立群と非自立群の間で有意差が認められた入院時変数

	Variables at admission
$n=120$	Age, Post-stroke time at admission, FIM [®] dressing, SIAS verticality, SIAS abdominal muscle strength, SIAS visuospatial perception, Berg balance scale, STEF affected side, STEF unaffected side, HDS-R
$n=100$	Post-stroke time at admission, FIM [®] dressing, SIAS verticality, SIAS abdominal muscle strength, SIAS visuospatial perception, Berg balance scale, STEF unaffected side, HDS-R
$n=80$	Age, Post-stroke time at admission, FIM [®] dressing, SIAS verticality, SIAS abdominal muscle strength, SIAS visuospatial perception, Berg balance scale, STEF affected side, STEF unaffected side, HDS-R
$n=60$	Age, Post-stroke time at admission, FIM [®] dressing, SIAS verticality, SIAS abdominal muscle strength, SIAS visuospatial perception, Berg balance scale, STEF affected side, STEF unaffected side, HDS-R
$n=40$	Age, Post-stroke time at admission, FIM [®] dressing, SIAS verticality, SIAS abdominal muscle strength, SIAS visuospatial perception, Berg balance scale, STEF affected side, STEF unaffected side, HDS-R

Abbreviations: SIAS, stroke impairment assessment set; HDS-R, revised Hasegawa's dementia scale; STEF, simple test for evaluating hand function.

表 3. 人工ニューラルネットワークとロジスティック回帰モデルの精度の比較

		Accuracy (%)	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)
$n=120$	ANN ^a	88.3	83.2	91.7	81.7	91.4
	LR ^a	80.0]*	72.7	85.4]*	65.8]*	86.7
$n=100$	ANN ^b	84.0	82.2	86.9	68.3	91.0
	LR ^b	75.0]*	65.8]*	80.5	56.7	83.6
$n=80$	ANN ^c	85.0	86.7	85.9	63.3	94.3
	LR ^c	73.8]*	69.7	82.8	56.7	82.7
$n=60$	ANN ^d	83.3	80.0	89.0	75.0	88.5
	LR ^d	80.0	73.3	87.5	70.0	86.0
$n=40$	ANN ^e	90.0	83.3	100.0	100.0	86.7
	LR ^e	82.5	85.4	88.3	70.0	90.0

* $p < 0.05$.

Abbreviations: ANN, artificial neural network; LR, logistic regression; PPV, positive-predictive value; NPV, negative-predictive value.

^aModels created by SIAS verticality, Berg balance scale, and STEF on unaffected side.

^bModels created by time post-stroke, Berg balance scale, and STEF on unaffected side.

^cModels created by Berg balance scale and Revised Hasegawa's dementia scale.

^dModels created by age, Berg balance scale, and STEF on unaffected side.

^eModels created by Berg balance scale, and STEF on unaffected side.

れなかったが、陽性的中率、陰性的中率、感度、特異度のすべてで人工ニューラルネットワークがロジスティック回帰を上回っていた。感度と特異度および陽性的中率と陰性的中率はそれぞれトレードオフの関係にあるため、その一方のみが高値であっても予測精度が高いことを意味しないが、100名のデータセットではこれらすべての項目で人工ニューラルネットワークが上回り、その一部の項目で有意差が認められていることから、われわれは100名のデータセットにおいても人工ニューラルネットワークのモデルがロジスティック回帰と比べて予測精度の高いモデルが作成されたと解釈している。一方、60名のデータセットでも、正分類率、陽性的中率、陰性的中率、感度、特異度のすべてで人工ニューラルネットワークが上回っているが、両者の数値は似通っており、有意な差が認められなかったため、われわれは両モデルの間に精度の

差はないと解釈した。

なお本研究では、脳卒中患者が一連の更衣を監視や介助なく遂行できるか否かを予測することが臨床上重要と考え、上半身と下半身の更衣自立度を包括的に扱った。しかし、今回の FIM[®] instrument の更衣上半身と下半身の成績をみると、上半身に比べて下半身のほうが若干得点が低かった。これは更衣全体を自立できるか否かは、より難易度の高い下半身の更衣を自立できるかが大きな影響力をもっていることを意味し、本研究の結果は下半身の更衣の自立度をより強く反映した結果である可能性がある。

本研究の知見は、リハビリテーション分野において単一施設で独自の更衣自立予測モデルを作成する際に有益な情報になると考えられる。つまり、単一施設で更衣自立の予測モデルを作成する際にサンプル数が80以上収集できた場合は予測精度を高めるためには

人工ニューラルネットワークの利用が望ましいと考えられ、サンプル数が60以下のケースであればロジスティック回帰でも人工ニューラルネットワークと同等の精度のモデルが期待できるといえる。今後の課題は、上半身と下半身の更衣で別々の分析を行って結果の相違の有無を検証すること、今回は更衣の自立可否を指標とした検証であったが別のADLでも同様の結果となるかを確認すること、他の施設のデータでも同様の結果が得られるかを確認することが必要な点である。またモデルの独立変数を変更した検討も必要である。

謝辞

本研究はJSPS 科研費18K17728の助成を受けたものです。

The use of the FIM[®] instrument to collect data for this research study was authorized and conducted in accordance with the term of a special purpose license granted to the licensee by Uniform Data System for Medical Rehabilitation (UDSMR). The licensee has not been trained by UDSMR in the use of the FIM[®] instrument, and the patient data collected during the course of this research study has not been submitted to or processed by UDSMR. No implication is intended that such data has been or will be subjected to the standard data processing procedures of UDSMR or that it is otherwise comparable with the data processed by UDSMR. FIM[®] is a trademark of UDSMR, a division of UB Foundation Activities, Inc.

文献

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. *Circulation* 2015; 313: e29–322.
2. Counsell C, Dennis M. Systematic review of prognostic models in patients with acute stroke. *Cerebrovasc Dis* 2001; 12: 159–70.
3. Veerbeek JM, Kwakkel G, van Wegen EE, Ket JC, Heymans MW. Early prediction of outcome of activities of daily living after stroke: a systematic review. *Stroke* 2011; 42: 1482–8.
4. Sato A, Fujita T, Ohashi Y, Yamamoto Y, Suzuki K, Otsuki K. A prediction model for activities of daily living for stroke patients in a convalescent rehabilitation ward. *J Allied Health Sci* 2016; 7: 1–6.
5. Fujita T, Sato A, Narita A, Sone T, Iokawa K, Tsuchiya K, et al. Use of a multilayer perceptron to create a prediction model for dressing independence in a small sample at a single facility. *J Phys Ther Sci* 2019; 31: 69–74.
6. Data Management Service of the Uniform Data System for Medical Rehabilitation and the Center for functional Assessment Research. Guide for the Uniform Data Set for Medical Rehabilitation, Version 3.0. New York: State University of New York at Buffalo; 1990.
7. Chino N, Tsubahara A, Sonoda S, Domen K, Takahashi H. Functional evaluation of stroke patients—SIAS and FIM (basic). 1st ed. Tokyo: Kanehara-shuppan; 2014. p. 78–138. Japanese.
8. Saito Y, Toshima M, Nori K, Kimura N. Factors affecting upper-body ability in post-acute stroke patients: a functional assessment. *Sagayouryouhou* 2012; 31: 134–40. Japanese.
9. Suzuki M, Omori M, Hatakeyama M, Yamada S, Matsushita K, Iijima S. Predicting recovery of upper-body dressing ability after stroke. *Arch Phys Med Rehabil* 2006; 87: 1496–502.
10. Fujita T, Sato A, Togashi Y, Kasahara R, Ohashi T, Yamamoto Y. Contribution of abdominal muscle strength for various activities of daily living in stroke patients with mild paralysis. *J Phys Ther Sci* 2015; 27: 815–8.
11. Fujita T, Nagayama H, Sato A, Yamamoto Y, Yamane K, Otsuki K, et al. Hierarchy of dysfunction related to dressing performance in stroke patients: a path analysis study. *PLoS ONE* 2016; 11: e0151162.
12. Fujita T, Sato A, Yamamoto Y, Yamane K, Otsuki K, Tsuchiya K, et al. Relationship between dressing and motor functions in stroke patients: a study with partial correlation analysis. *J Phys Ther Sci* 2015; 27: 3771–4.
13. Chino N, Tsubahara A, Sonoda S, Domen K, Takahashi H. Functional evaluation of stroke patients—SIAS and FIM (basic). Tokyo: Kanehara-shuppan; 2012. p. 36–67. Japanese.
14. Berg K, Wood-Dauphinee S, Williams JJ, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can* 1989; 41: 304–11.
15. Kaneko T, Muraki T. Development and standardization of the hand function test. *Bull Allied Med Sci Kobe* 1990; 6: 49–54.
16. Imai Y, Hasegawa K. The revised Hasegawa's dementia scale [HDS-R]: evaluation of its usefulness as a screening test for dementia. *J Hong Kong Coll Psychiatr* 1994; 4: 20–4.
17. Ottenbacher KJ, Smith PM, Illig SB, Linn RT, Fiedler RC, Granger CV. Comparison of logistic regression and neural networks to predict rehospitalization in patients with stroke. *J Clin Epidemiol* 2001; 54: 1159–65.
18. Ottenbacher KJ, Linn RT, Smith PM, Illig SB, Mancuso M, Granger CV. Comparison of logistic regression and neural network analysis applied to predicting living setting after hip fracture. *Ann Epidemiol* 2004; 14: 551–9.
19. Heydari ST, Ayatollahi SM, Zare N. Comparison of artificial neural networks with logistic regression for detection of obesity. *J Med Syst* 2012; 36: 2449–54.
20. Tang ZH, Liu J, Zeng F, Li Z, Yu X, Zhou L. Comparison of prediction model for cardiovascular autonomic dysfunction using artificial neural network and logistic regression analysis. *PLoS One* 2013; 8: e70571.
21. Mendes RG, de Souza CR, Machado MN, Correa PR, Di Thommazo-Luporini L, Arena R, et al. Predicting reintubation, prolonged mechanical ventilation and death in post-coronary artery bypass graft surgery: a comparison between artificial neural networks and logistic regression models. *Arch Med Sci* 2015; 11: 756–63.
22. Linder R, König IR, Weimar C, Diener HC, Pöppel SJ, Ziegler A. Two models for outcome prediction — a

- comparison of logistic regression and neural networks. *Methods Inf Med* 2006; 45: 536–40.
23. Lundin M, Lundin J, Burke HB, Toikkanen S, Pylkkänen L, Joensuu H. Artificial neural networks applied to survival prediction in breast cancer. *Oncology* 1999; 57: 281–6.
 24. Kim WO, Kil HK, Kang JW, Park HR. Prediction on lengths of stay in the postanesthesia care unit following general anesthesia: preliminary study of the neural network and logistic regression modelling. *J Korean Med Sci* 2000; 15: 25–30.
 25. Lin SP, Lee CH, Lu YS, Hsu LN. A comparison of MICU survival prediction using the logistic regression model and artificial neural network model. *J Nurs Res* 2006; 14: 306–14.
 26. Liew PL, Lee YC, Lin YC, Lee TS, Lee WJ, Wang W, et al. Comparison of artificial neural networks with logistic regression in prediction of gallbladder disease among obese patients. *Dig Liver Dis* 2007; 39: 356–62.
 27. Lin CC, Ou YK, Chen SH, Liu YC, Lin J. Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture. *Injury* 2010; 41: 869–73.
 28. Shi HY, Lee KT, Lee HH, Ho WH, Sun DP, Wang JJ, et al. Comparison of artificial neural network and logistic regression models for predicting in-hospital mortality after primary liver cancer surgery. *PLoS One* 2012; 7: e35781.
 29. Parsaeian M, Mohammad K, Mahmoudi M, Zeraati H. Comparison of logistic regression and artificial neural network in low back pain prediction: second national health survey. *Iran J Public Health* 2012; 41: 86–92.
 30. Faradmal J, Soltanian AR, Roshanaei G, Khodabakhshi R, Kasaeian A. Comparison of the performance of log-logistic regression and artificial neural networks for predicting breast cancer relapse. *Asian Pac J Cancer Prev* 2014; 15: 5883–8.
 31. Ozden S, Er S, Saylam B, Yildiz BD, Senol K, Tez M. A comparison of logistic regression and artificial neural networks in predicting central lymph node metastases in papillary thyroid microcarcinoma. *Ann Ital Chir* 2018; 89: 193–8.
 32. Hassanipour S, Ghaem H, Arab-Zozani M, Seif M, Fararouei M, Abdzadeh E, et al. Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: a systematic review and meta-analysis. *Injury* 2019; 50: 244–50.