

Artificial neural network models for predicting 1-year mortality in elderly patients with intertrochanteric fractures in China

L. Shi¹, X.C. Wang¹ and Y.S. Wang²

¹Information Science and Technology College, Dalian Maritime University, Dalian, China

²Department of Orthopedics, First Affiliated Hospital, Liaoning Medical University, Jinzhou, China

Abstract

The mortality rate of older patients with intertrochanteric fractures has been increasing with the aging of populations in China. The purpose of this study was: 1) to develop an artificial neural network (ANN) using clinical information to predict the 1-year mortality of elderly patients with intertrochanteric fractures, and 2) to compare the ANN's predictive ability with that of logistic regression models. The ANN model was tested against actual outcomes of an intertrochanteric femoral fracture database in China. The ANN model was generated with eight clinical inputs and a single output. ANN's performance was compared with a logistic regression model created with the same inputs in terms of accuracy, sensitivity, specificity, and discriminability. The study population was composed of 2150 patients (679 males and 1471 females): 1432 in the training group and 718 new patients in the testing group. The ANN model that had eight neurons in the hidden layer had the highest accuracies among the four ANN models: 92.46 and 85.79% in both training and testing datasets, respectively. The areas under the receiver operating characteristic curves of the automatically selected ANN model for both datasets were 0.901 (95%CI = 0.814-0.988) and 0.869 (95%CI = 0.748-0.990), higher than the 0.745 (95%CI = 0.612-0.879) and 0.728 (95%CI = 0.595-0.862) of the logistic regression model. The ANN model can be used for predicting 1-year mortality in elderly patients with intertrochanteric fractures. It outperformed a logistic regression on multiple performance measures when given the same variables.

Key words: Artificial neural network; Intertrochanteric fracture; Outcome prediction; One-year mortality

Introduction

In recent years, the incidence of hip fractures has been increasing with the aging of populations worldwide, and approximately 50% of hip fractures are intertrochanteric fractures (1). In China, the percentage of the population over the age of 65 was 5% in 1982 and now stands at 7.5% but is expected to rise to more than 15% by 2025 (2). Families have been getting smaller and fracture care capabilities are declining because of the one-child family policy. The literature reports that there is an increased risk of death after intertrochanteric fracture, with 1-year mortality ranging from 8.4 to 36%, which imposes a tremendous economic burden on health care (3).

As most of the frailty syndrome that commonly occurs after hip fractures are known risk factors for mortality, it is necessary to identify those patients who are candidates

for interventions, in order to reduce their risk for mortality (4,5). Accurately predicting mortality in elderly patients with intertrochanteric fractures is a significant clinical challenge, which is essential for family counseling and for designing a personalized rehabilitation program (6,7).

An artificial neural network (ANN) is a mathematical model that is inspired by the structure and/or functional aspects of biological neural networks (8). Most of the time, an ANN is an adaptive system that makes new decisions, classifications, and forecasts based on external or internal information that flows through the network during the learning phase (9). ANNs have been successfully applied in many complex and diverse tasks in clinical medicine, such as clinical outcome predictions of head injury (10), bone mineral density (11), and living settings, especially in predicting hip fracture mortality (12,13).

Correspondence: Y.S. Wang, Department of Orthopedics, First Affiliated Hospital, Liaoning Medical University, No. 5-2 Renmin Street, Guta District, Jinzhou 121000, Liaoning Province, China. E-mail: wys200910130@163.com

Received January 16, 2013. Accepted August 7, 2013. First published online November 18, 2013.

Few published reports, however, have comprehensively assessed mortality in elderly patients with intertrochanteric fractures. We hypothesized that the ability to identify 1-year mortality after intertrochanteric fracture could be improved by using computer analyses involving neural networks. To test this hypothesis, we trained an ANN by using a provincial database in China, and tested the network against a logistic regression model for comparison.

Patients and Methods

Study design and patient population

A total of 2769 patients treated for intertrochanteric femoral fractures between August 2001 and June 2010 were enrolled in this study. Data were obtained from the Department of Orthopedics, First Affiliated Hospital, Liaoning Medical University. Two hundred and seven patients were excluded due to age <65 years (93 patients), pathological fractures (32 patients), ineligibility for surgery (58 patients), and multiple fractures (24 patients). Information about follow-up for ≥ 1 year or death within 1 year of surgery was obtained for 2150 of the 2562 patients (Figure 1).

Eight predictor variables - age, gender, nursing home, the New Mobility Score (NMS), dementia or cognitive impairment, diabetes, cancer, and cardiac disease - were chosen to build a prediction model for 1-year survival because of their previously established influence on

patient outcomes after intertrochanteric fracture surgery (12-15). Dementia/cognitive impairment was determined by the Clinical Dementia Rating scale. Cardiac disease was recorded from the medical record as previously known clinical evidence of heart disease confirmed by a typical history, positive results of an ECG, or exercise testing. The 1-year mortality risk was the primary outcome (Figure 2). The data and its format were similar for both ANN and logistic regression models.

This study was carried out in accordance with the ethics standards in the updated version of the 1964 Declaration of Helsinki. The study protocol was approved by the Local Ethics Committees of all participating centers, and the patients gave informed, oral and written consent before participating in the study.

Performance and accuracy

Logistic regression modeling. As a generalized linear regression model, logistic regression is widely used for predicting the probability of occurrence of an event. In this study, the dataset was divided randomly into two parts: 1432 cases for training and 718 cases for testing the model. The model was built by using a training set with logistic regression. Age, gender, nursing home, the NMS, dementia or cognitive impairment, diabetes, cancer, and cardiac disease were the independent variables, and the 1-year survival was the outcome. The logistic regression data were analyzed using Intercooled STATA 8.0 for Windows (StataCorp LP, USA). We used the test dataset

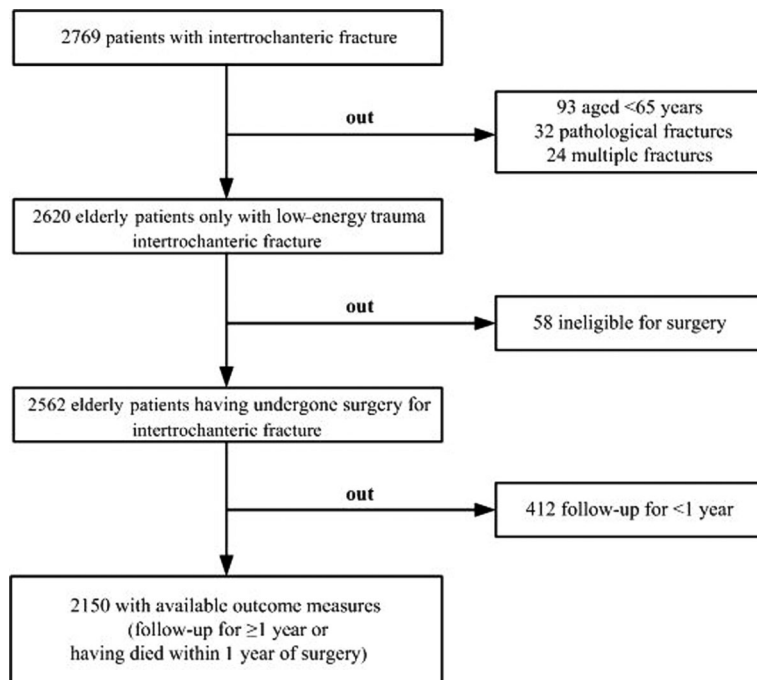


Figure 1. Stepwise selection of patients.

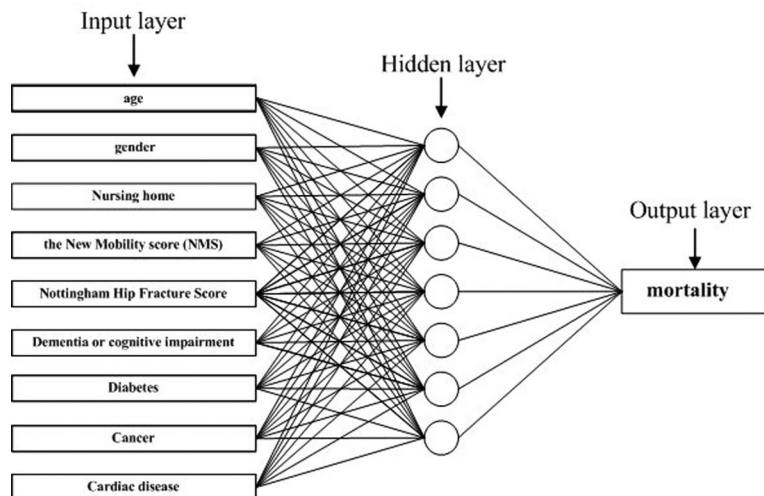


Figure 2. Schematic representation showing the structure of the artificial neural network models, which have 8 input nodes, 6 nodes in hidden layer, and 1 output node, which represents 1-year mortality in elderly patients with intertrochanteric fracture.

(718 cases) to test the logistic model we built.

ANN modeling

A total of 2150 patients were randomly assigned to training (1432 cases; 66.6%) and testing (718 cases; 33.4%) datasets. A schematic representation of our model is shown in Figure 2, and training in our model relied on a technique called “informative sampling” that we have described in a feed-forward three-layer neural network. The input layer consisted of eight input nodes, the hidden layer consisted of six hidden nodes, and the output layer consisted of one output node (16).

The learning rate for network training was set to 0.20. A trial and error procedure was employed to optimize the number of network layers, hidden neurons, and stopping criteria. Training consisted of pairing the eight inputs with the known output and allowing the algorithm to adjust the weight of the connections for different variables. The strength of these connections changes with repeated exposure to data and is rectified by feedback (10). We trained 20 ANNs simultaneously, which allowed us to produce a single prediction by averaging the 20 individual predictions. In this way, the four most accurate ANNs were identified for analysis, and two logistic regression models were evaluated against the ANNs using the same datasets for each training and testing dataset.

Comparison of performance by the two models

Continuous variables are reported as means \pm SD. Means were compared using the *t*-test. Categorical variables are reported as counts and percentages, and the chi-square test was used to detect associations between variables. Statistical analyses were performed using PASW Statistics 13.0 and Clementine 11.0 softwares (SPSS Inc., USA). The significance level was set at 0.05. Prediction accuracy was evaluated by comparing the logistic regression and ANN models with a set of

patients randomly selected from the database not previously exposed to the models. The overall accuracy $[(\text{true positives} + \text{true negatives}) / (\text{true positives} + \text{true negatives} + \text{false positives} + \text{false negatives})]$ of the final model was determined by comparing the predicted values with the actual events. The receiver operating characteristic (ROC) curves were constructed by plotting true positives (i.e., sensitivity) vs the false-positive fraction (i.e., one minus specificity) (17).

Results

A total of 679 males and 1471 females were included in the study. The average age at the time of surgery was 81.6 years, ranging from 65 to 98 years. During surgery, dynamic hip screws were used in femoral intertrochanteric fractures; 582 patients died within 1 year of surgery (27.0%), and the other 1568 (73.0%) survived for ≥ 1 year with no patients lost to follow-up. Table 1 shows the clinical characteristics of the training, testing, and group lost to follow-up datasets. These three datasets did not differ significantly for any variable ($P > 0.05$).

The four most accurate ANNs were pooled and compared with the two different regression models for predicting 1-year mortality of elderly patients with intertrochanteric fractures in China and are presented in Table 2. We specified 10, 15, and 20 neurons in the hidden layer for the ANN models and built a model automatically selected by the computer (8 neurons in the hidden layer) (13). The ANN model that had eight neurons in the hidden layer had the highest accuracies of 92.46 and 85.79% in both training and testing datasets among the four ANN models, respectively. One logistic model contained the main effects, while the other model added two-way interactions. The logistic model contained two-way interactions with higher accuracies in both the training (80.92 vs 72.08%) and testing (71.33 vs 67.59%) datasets.

Table 1. Description of the baseline variables within the training, testing, and lost to follow-up datasets.

	Training (n = 1432)	Testing (n = 718)	Follow-up for <1 year (n = 412)	Total (n = 2150)
Age				
<80 years	433 (30.2%)	196 (27.3%)	119 (28.9%)	629 (29.3%)
≥80 years	999 (69.8%)	522 (72.7%)	293 (71.1%)	1521 (70.7%)
Gender				
Female	982 (68.6%)	491 (68.4%)	285 (69.2%)	1473 (68.5%)
Male	450 (31.4%)	227 (31.6%)	127 (30.8%)	677 (31.5%)
Nursing home				
Poor	328 (22.9%)	166 (23.1%)	87 (21.1%)	494 (23.0%)
Good	1104 (77.1%)	552 (76.9%)	325 (78.9%)	1656 (77.0%)
New Mobility Score				
Mobility score ≥5	963 (67.2%)	490 (68.2%)	287 (69.7%)	1453 (67.6%)
Mobility score <5	469 (32.8%)	228 (31.8%)	125 (30.3%)	697 (32.4%)
Dementia or cognitive impairment				
No	1107 (77.3%)	562 (78.3%)	313 (76.0%)	1669 (77.6%)
Yes	325 (22.7%)	156 (21.7%)	99 (24.0%)	481 (22.4%)
Diabetes				
No	939 (65.6%)	476 (66.3%)	260 (63.1%)	1415 (65.8%)
Yes	493 (34.4%)	242 (33.7%)	152 (36.9%)	735 (34.2%)
Cancer				
No	1218 (85.1%)	572 (79.7%)	354 (85.9%)	1790 (83.3%)
Yes	214 (14.9%)	146 (20.3%)	58 (14.1%)	360 (16.7%)
Cardiac disease				
No	972 (67.9%)	445 (62.0%)	288 (69.9%)	1417 (65.9%)
Yes	460 (32.1%)	273 (38.0%)	124 (30.1%)	733 (34.1%)
Death				
No	1051 (73.4%)	518 (72.1%)	337 (81.8%)	1569 (73.0%)
Yes	381 (26.6%)	200 (27.9%)	75 (18.2%)	581 (27.0%)

There were no significant differences between variables (*t*-test).

The automatically selected ANN model (8 neurons) and the logistic model that included main effects and two-way interactions had the highest accuracies. The areas under the ROC curves of the ANN model were higher than those of the logistic regression model for training and

testing datasets. The areas under the ROC curves of the automatically selected ANN model for both datasets were 0.901 (95%CI=0.814-0.988) and 0.869 (95%CI=0.748-0.990), higher than the 0.745 (95%CI=0.612-0.879) and 0.728 (95%CI=0.595-0.862) of the logistic regression

Table 2. Accuracies of artificial neural network (ANN) models and logistic models.

Model	Accuracy	
	Training (n = 1432)	Testing (n = 718)
ANN		
10 neurons in the hidden layer	86.32%	83.20%
15 neurons in the hidden layer	81.53%	74.66%
20 neurons in the hidden layer	84.14%	72.85%
Automatically selected ^a	92.46%	85.79%
Logistic models		
Main effects	72.08%	67.59%
Main effects + two-way interactions	80.92%	71.33%

^aAutomatically selected ANN model by the software, with 8 neurons in the hidden layer.

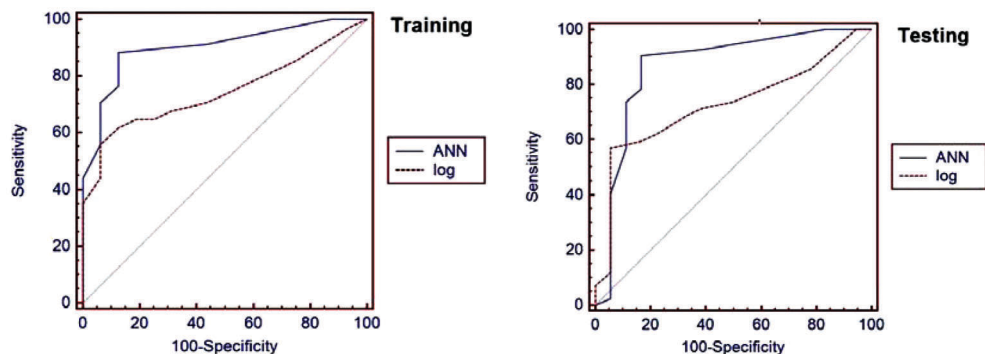


Figure 3. Comparison of ROC curves for artificial neural network (ANN) and logistic regression analysis in both training and testing dataset.

model (Figure 3).

Discussion

Osteoporotic intertrochanteric fractures in older adults are complicated by the many variables affecting treatment outcomes in elderly patients, including age, fracture type, preexisting comorbidities, and prefracture mobility status (18-20). The outcome of the 1-year mortality rate in elderly patients with hip fractures after surgical treatment is about 30% (12,13). Research suggests that the incidence and cost of treating intertrochanteric fractures in China are rising with the aging population (21). The idea of prediction models based on related clinical data at admission is worth trying and seems to be of practical value in this situation. Few studies, however, use clinical variables to provide an accurate prediction model for the 1-year mortality rate after surgery in elderly patients with intertrochanteric fractures in China. The development of such a model could provide doctors, patients, and their families with more objective information and design a personalized rehabilitation program in the future.

In this study, a clinical database was used to train and test a neural network in predicting outcomes for the 1-year mortality of elderly patients with intertrochanteric fractures. The neural network performance was compared with two logistic regression models created with the same clinical input and outcome measures. In this study, the automatically selected ANN model (8 neurons) had the highest accuracies of 92.46 and 85.79% in both training and testing datasets among the four ANN models. The average prediction rate obtained by ANN was 89.22% (training group: 92.46% vs testing group: 85.79%), while the average prediction rate obtained by logistic regression was 76.13% (training group: 80.92% vs testing group: 71.33%). The areas under the ROC curves of the automatically selected ANN model for both datasets were 0.901 (95%CI=0.814-0.988) and 0.869 (95%CI=0.748-0.990), higher than the 0.745 (95%CI=0.612-0.879) and 0.728 (95%CI=0.595-0.862) values with the logistic

regression model. We found that ANN performance surpasses that of logistic regression models by using the same limited variables, which are similar to those obtained by Ottenbacher et al. (12) and Lin et al. (13) in predicting outcomes in elderly patients with hip fractures after surgery.

ANN as a form of modeling has been sufficiently utilized in the field of clinical outcome prediction (22-26). Some studies have focused on the role of ANN in predicting the mortality of hip fractures (12,13), but research focused on intertrochanteric fractures is rare. The number of inputs in our ANN model was limited, but the result is better than with logistic regression models regarding accuracy and specificity when using the same test set. Neural networks are significantly better than logistic regression models in this research, perhaps because they are not affected by interactions between factors. If the number of model inputs is added later and the predictions made are more specific, this discrepancy increases further.

Our ANN model suffers from several limitations. To rapidly apply the predictive models clinically, we simplified or quantified some variables artificially, such as age and the NMS (27,28). The inputs do not include implants, detailed types of fractures, and different operating methods because of the database itself (15,29). Logistic regression models with interaction analyses were not fully compared to ANN models. This strategy may potentially present better accuracy levels for logistic regression models. Accurately predicting outcomes for intertrochanteric fractures with ANN models will ultimately require predicting more than just 1-year mortality, which requires large datasets with detailed clinical as well as long-term follow-up information. We are now building larger datasets that we can use to predict more sophisticated fracture data than is possible with the current model.

Acknowledgments

The authors would like to express their gratitude to Yang Li at the Liaoning Medical University for her editorial help.

References

- Bhandari M, Schemitsch E, Jonsson A, Zlowodzki M, Haidukewych GJ. Gamma nails revisited: gamma nails versus compression hip screws in the management of intertrochanteric fractures of the hip: a meta-analysis. *J Orthop Trauma* 2009; 23: 460-464, doi: 10.1097/BOT.0b013e318162f67f.
- Hesketh T, Lu L, Xing ZW. The effect of China's one-child family policy after 25 years. *N Engl J Med* 2005; 353: 1171-1176, doi: 10.1056/NEJMhr051833.
- Abrahamsen B, van Staa T, Ariely R, Olson M, Cooper C. Excess mortality following hip fracture: a systematic epidemiological review. *Osteoporos Int* 2009; 20: 1633-1650, doi: 10.1007/s00198-009-0920-3.
- Cree M, Soskolne CL, Belseck E, Hornig J, McElhaney JE, Brant R, et al. Mortality and institutionalization following hip fracture. *J Am Geriatr Soc* 2000; 48: 283-288.
- Ensrud KE, Ewing SK, Taylor BC, Fink HA, Stone KL, Cauley JA, et al. Frailty and risk of falls, fracture, and mortality in older women: the study of osteoporotic fractures. *J Gerontol A Biol Sci Med Sci* 2007; 62: 744-751, doi: 10.1093/gerona/62.7.744.
- Hagino T, Ochiai S, Sato E, Watanabe Y, Senga S, Haro H. Prognostic prediction in patients with hip fracture: risk factors predicting difficulties with discharge to own home. *J Orthop Traumatol* 2011; 12: 77-80, doi: 10.1007/s10195-011-0138-y.
- Hirose J, Ide J, Yakushiji T, Abe Y, Nishida K, Maeda S, et al. Prediction of postoperative ambulatory status 1 year after hip fracture surgery. *Arch Phys Med Rehabil* 2010; 91: 67-72, doi: 10.1016/j.apmr.2009.09.018.
- Greenwood D. An overview of neural networks. *Behav Sci* 1991; 36: 1-33, doi: 10.1002/bs.3830360102.
- Terrin N, Schmid CH, Griffith JL, D'Agostino RB, Selker HP. External validity of predictive models: a comparison of logistic regression, classification trees, and neural networks. *J Clin Epidemiol* 2003; 56: 721-729, doi: 10.1016/S0895-4356(03)00120-3.
- Rughani AI, Dumont TM, Lu Z, Bongard J, Horgan MA, Penar PL, et al. Use of an artificial neural network to predict head injury outcome. *J Neurosurg* 2010; 113: 585-590, doi: 10.3171/2009.11.JNS09857.
- Ongphiphadhanakul B, Rajatanavin R, Chailurkit L, Piaseu N, Teerarungsikul K, Sirisriro R, et al. Prediction of low bone mineral density in postmenopausal women by artificial neural network model compared to logistic regression model. *J Med Assoc Thai* 1997; 80: 508-515.
- 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: 551-559, doi: 10.1016/j.annepidem.2003.10.005.
- 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: 869-873, doi: 10.1016/j.injury.2010.04.023.
- Hu F, Jiang C, Shen J, Tang P, Wang Y. Preoperative predictors for mortality following hip fracture surgery: a systematic review and meta-analysis. *Injury* 2012; 43: 676-685, doi: 10.1016/j.injury.2011.05.017.
- Geiger F, Zimmermann-Stenzel M, Heisel C, Lehner B, Daecke W. Trochanteric fractures in the elderly: the influence of primary hip arthroplasty on 1-year mortality. *Arch Orthop Trauma Surg* 2007; 127: 959-966, doi: 10.1007/s00402-007-0423-7.
- Landwehr N, Hall M, Frank E. Logistic model trees. *Machine Learning* 2005; 59: 161-205, doi: 10.1007/s10994-005-0466-3.
- Eftekhari B, Mohammad K, Ardebili HE, Ghodsi M, Ketabchi E. Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. *BMC Med Inform Decis Mak* 2005; 5: 3, doi: 10.1186/1472-6947-5-3.
- Holvik K, Ranhoff AH, Martinsen MI, Solheim LF. Predictors of mortality in older hip fracture inpatients admitted to an orthogeriatric unit in Oslo, Norway. *J Aging Health* 2010; 22: 1114-1131, doi: 10.1177/0898264310378040.
- Jamal SY, Umer M, Khan A, Ullah Khan NA. Functional outcome, mortality and in-hospital complications of operative treatment in elderly patients with hip fractures in the developing world. *Int Orthop* 2010; 34: 431-435, doi: 10.1007/s00264-009-0803-4.
- Roche JJ, Wenn RT, Sahota O, Moran CG. Effect of comorbidities and postoperative complications on mortality after hip fracture in elderly people: prospective observational cohort study. *BMJ* 2005; 331: 1374, doi: 10.1136/bmj.38643.663843.55.
- Dai K, Zhang Q, Fan T, Sen SS. Estimation of resource utilization associated with osteoporotic hip fracture and level of post-acute care in China. *Curr Med Res Opin* 2007; 23: 2937-2943, doi: 10.1185/030079907X233061.
- Miyajima F, Furlan JC, Aarabi B, Arnold PM, Fehlings MG. Acute cervical traumatic spinal cord injury: MR imaging findings correlated with neurologic outcome - prospective study with 100 consecutive patients. *Radiology* 2007; 243: 820-827, doi: 10.1148/radiol.2433060583.
- Kaczmarczyk K, Wit A, Krawczyk M, Zaborski J, Gajewski J. Associations between gait patterns, brain lesion factors and functional recovery in stroke patients. *Gait Posture* 2012; 35: 214-217, doi: 10.1016/j.gaitpost.2011.09.009.
- Chang CL, Jin Z, Chang HC, Cheng AC. From neuromuscular activation to end-point locomotion: An artificial neural network-based technique for neural prostheses. *J Biomech* 2009; 42: 982-988, doi: 10.1016/j.jbiomech.2009.03.030.
- 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: 1159-1165, doi: 10.1016/S0895-4356(01)00395-X.
- Jaimes F, Farbiarz J, Alvarez D, Martinez C. Comparison between logistic regression and neural networks to predict death in patients with suspected sepsis in the emergency room. *Crit Care* 2005; 9: R150-R156, doi: 10.1186/cc3054.
- Kristensen MT, Foss NB, Ekdahl C, Kehlet H. Prefracture functional level evaluated by the New Mobility Score predicts in-hospital outcome after hip fracture surgery. *Acta Orthop* 2010; 81: 296-302, doi: 10.3109/17453674.2010.487240.
- Fan L, Dang X, Wang K. Comparison between bipolar

- hemiarthroplasty and total hip arthroplasty for unstable intertrochanteric fractures in elderly osteoporotic patients. *PLoS One* 2012; 7: e39531, doi: 10.1371/journal.pone.0039531.
29. Madsen JE, Naess L, Aune AK, Alho A, Ekeland A, Stromsoe K. Dynamic hip screw with trochanteric stabilizing plate in the treatment of unstable proximal femoral fractures: a comparative study with the Gamma nail and compression hip screw. *J Orthop Trauma* 1998; 12: 241-248, doi: 10.1097/00005131-199805000-00005.