Research Article

A Neural Network Approach to Predict Acute Allograft Rejection in Liver Transplant Recipients Using Routine Laboratory Data

Abdolhossein Zare,[1] Mohammad Amin Zare,[2] Neda Zarei,[1] Ramin Yaghoobi,[1] Mohammad Ali Zare,[1] Saeede Salehi,[3] Bita Geramizadeh,[1] Seid Ali Malekhosseini,[1] and Negar Azarpira[1,4,*]

[1]Transplant Research Center, Namazi Hospital, Shiraz University of Medical Sciences, Shiraz, IR Iran
[2]School of Electrical and Computer Engineering, Shiraz University, Shiraz, IR Iran
[3]Department of Medical Biotechnology, School of Paramedical Sciences, Shiraz University of Medical Sciences, Shiraz, IR Iran
[4]Shiraz Institute for Stem Cell and Regenerative Medicine, Shiraz, IR Iran

[*]Corresponding author: Negar Azarpira, Transplant Research Center, Namazi Hospital, Shiraz University of Medical Sciences, Shiraz, IR Iran. Tel: +98-7136473954, Fax: +98-7136473954, E-mail: negarazarpira@yahoo.com

Abstract

Background: Discovery of non-invasive methods for acute rejection in liver transplant patients would contribute to preservation of liver function in the graft. Recently, however, outcome prediction based on biostatistical models like artificial neural networks (ANNs) is increasingly becoming impressive in medicine.

Objectives: The aim of this study was to obtain a predictive model based on ANN technique and to figure out the best time for early prediction of acute allograft rejection after transplantation in liver transplant recipients.

Methods: Feed-forward, back-propagation neural network was developed to predict acute rejection in liver transplant recipients using clinical and biochemical data from 148 liver transplant recipients over days 3, 7, and 14 post-transplantation. Sensitivity and receiver-operating characteristic (ROC) analysis were done to reveal the importance of input variables and the performance of the neural network.

Results: The results were compared with a logistic regression (LR) model using the same data. Our results showed that the data related to day 7 gave the best results in terms of ANN performance; and the most important factors in the predictive model were aspartate aminotransferase (AST) and alanine aminotransferase (ALT). The ANN's accuracy was 90%, sensitivity was 87%, specificity was 90% in the testing set, and the performance of the ANN was better than that of the LR model. The ANN recognized correctly eight out of ten acute rejection patients and 34 out of 36 non-rejection ones in the testing set.

Conclusions: This study suggests that ANN could be a valuable adjunct to conventional liver function tests for monitoring liver transplant recipients in the early postoperative period.

Keywords: Liver Transplantation, Acute Rejection, Artificial Neural Networks, Logistic Regression

1. Background

Liver transplantation is an established therapeutic option for patients with end-stage chronic liver disease (1-4). Despite the development and improvement of immunosuppressive regimens and surgical techniques in the last decades, acute rejection (AR) still remains of fundamental problems in 10% - 20% of liver transplant (LT) patients and it is more common in the first few weeks post-transplantation. AR episodes are distinguished in 34% to 70% of patients, and 5% to 20% of patients will result in chronic rejection (CR), which is usually irreversible and needs re-transplantation (5-9). Currently, liver biopsy is a widely used gold standard for the examination of the rejection in liver transplant patients. Nevertheless, being invasive, difficult, and costly on one hand and accompa-

nying by significant morbidity and low mortality on the other hand are the drawbacks of this method. Thus, presenting a non-invasive, accurate, simple, and inexpensive method for detecting AR following LT can be a useful adjunct (6, 10-12). Recent studies have shown that artificial neural network (ANN) analysis is potentially more reliable than traditional statistical models are in predicting clinical outcomes (10, 12, 13). The authority of ANN is being generic in structure, being able to learn from historical data (14). Interestingly, ANNs are computer systems inspired by human neural physiology that comprise a set of highly interconnected processing units (neurons) tied with weighted connections, and contain an input layer, an output layer, and one or more hidden layers. The input and output layers are formed by different data available

for the analysis (e.g., various laboratory tests) and different outcomes, respectively ([15]). Learning through examples is one of the main features of ANN, which is achieved through exposure to paired input–output data (training). ANN learns to associate each of the inputs with the corresponding output by adjusting the weights of the connections between its neurons. Based on the knowledge collected during training, ANN assigns outputs to new inputs. As a result, once trained, the ANN can recognize patterns or make predictions on those data that are not selected for training or are unknown to the ANN ([12], [16]). ANNs as a data-modeling method have been studied in different areas of medical applications ([3], [13], [14], [17-20]). In the transplantation field, the use of ANNs has been considered in the prediction of fibrosis in hepatitis C virus infected liver transplant recipients ([12]), prediction of outcomes after liver transplantation ([21]), and prediction of graft failure ([22]). Hence, liver transplantation is a field that would largely benefit from such models. The present cross-sectional study was performed to assess the ability of ANN to predict AR in liver transplant recipients using clinical and biochemical data of the patients, and to find the most impressive post-transplantation day in terms of prediction of AR to create an ANN model for future studies.

2. Methods

2.1. Patients

Clinical and biochemical data were collected for 3rd, 7th, and 14th days post-transplantation from 210 patients transplanted between 2009 and 2015, of which 150 were non-rejection and 60 were acute rejection patients. Acute rejection was diagnosed histologically according to the established criteria ([23]). This study was conducted at Namazi Hospital affiliated to Shiraz University of Medical Sciences. Patients with missing data were excluded from the study. Therefore, finally a total of 148 patients were included on days 3, 7, and 14. The patients were randomly divided into two groups for each three days. The first group included the data of 24 acute rejection patients and 78 non-rejection patients. This group was used to develop trained ANN and LR models. The second group, comprising the rest of acute rejection patients (10 persons) and non-rejection patients (36 persons), was used to test the models.

2.2. Development of the Artificial Neural Networks

The neural network was developed using Python programming language V3.5.2 (https://www.python.org/) in Linux Ubuntu 16.04 LTS (https://www.ubuntu.com/) to predict acute rejection in liver transplant recipients. A feed-forward artificial neural network (FFANN) topology, also known as multilayer perceptron (MLP), was used with back propagation (BP) algorithm to build the predictive models. MLP is a neural network that has at least three layers of neurons: one input layer, one output layer, and at least one hidden layer. Neurons are tied together with weighted connections. The weighted output is then passed through a transfer function ([3], [15], [24]). In this study, the default transfer functions for hidden layers and output layer were tansig and purelin, respectively. Furthermore, of the 21 input variables, the following seven statistically significant variables retained for constructing ANN models: aspartate aminotransferase (IU/L), alanine aminotransferase (IU/l), alkaline phosphatase (IU/L), total serum bilirubin (mg/dL), direct serum bilirubin (mg/dL), total serum protein (g/dL), and serum calcium (mg/dL). We found one hidden layer with 10 neurons and two output neurons (rejection and non-rejection) ([Figure 1]). We preprocessed the input and target values: mapping them into the interval [-1, 1]. This simplifies the challenge for the network. It also ensures that targets fall into the range that your new feed-forward network can reproduce. Over-fitting is one major challenge of neural networks, meaning generating networks that are too closely adapted to the training data. One way to overcome this obstacle is to randomly apportion the data into a training set and a validation set as described above and to test the network response to validation set. In our study, the network performed accurate response to the validation set (testing set). Thus, it could be claimed that over-fitting has not happened in the training phase. The ANN was trained 10 times with 7-fold cross validation. At the end of each training session, the network was tested and the accuracy of prediction was calculated on a test group of 46 patients who were not chosen for training and whose outcome was unknown to the ANN. Then, the best network in terms of accuracy was selected.

2.3. Development of Logistic Regression Model

The logistic regression, a generalized linear model, was developed using Python programming language V3.5.2 in Linux Ubuntu 16.04 LTS. The training and testing datasets were the same as those applied in ANN; therefore, there were an ANN and a logistic model for training and testing datasets. The logistic regression performance was evaluated with ROC analysis using the same testing set.

2.4. Sensitivity Analysis on Artificial Neural Network Output

One of the major drawbacks of neural network models is being "black box" because of their failure to explain results or inability to generate interpretable parameters for each input variable ([12], [25]). To cope with this problem, sensitivity analysis was applied in order to investigate the importance of each input variable to the output.

Figure 1. Architecture of the Neural Networks Used in the Present Study
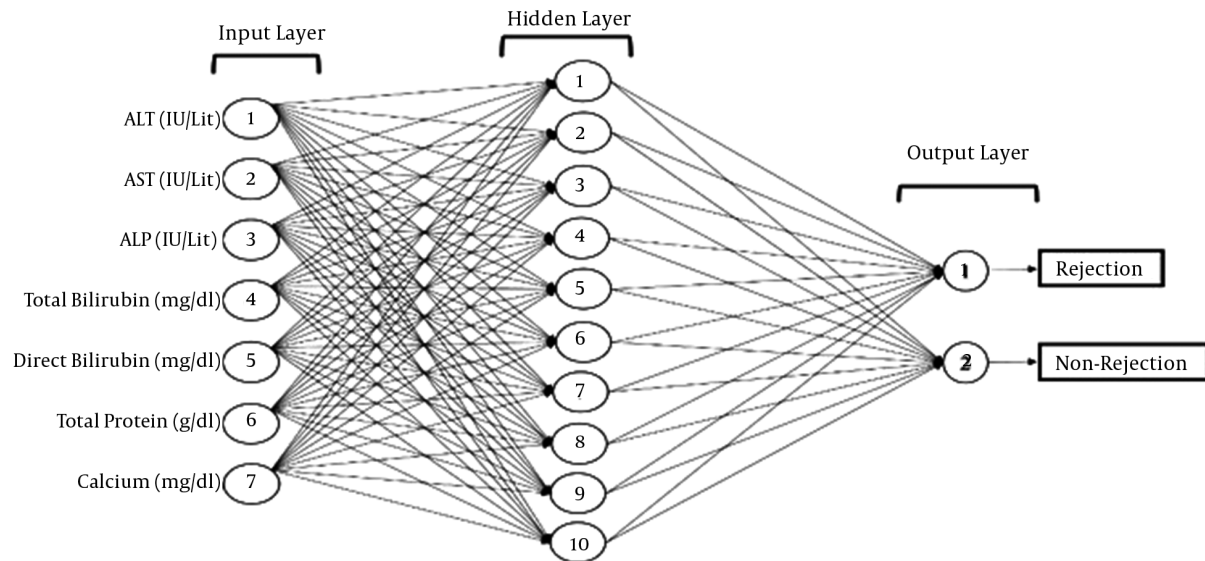
The network consists of three layers, with 7 neurons in the input layer, 10 neurons in the hidden layer, and 2 neurons in the output layer. AST, aspartate aminotransferase; ALP, alkaline phosphatase; ALT, alanine aminotransferase.

2.5. Statistical Analysis

All statistical analyses were performed using SPSS software (SPSS: An IBM Company, version 19.0, IBM Corporation, Armonk, NY, USA). Continuous variables expressed as mean (standard deviation (SD)) were compared using independent t-test. The Chi-square test was used for categorical variables. Receiver-operating characteristic (ROC) curve analysis was performed to assess the performance or accuracy of the trained neural networks (26). The predictive performance was evaluated by calculating true-positive and false-positive fractions, positive and negative predictive values (PPV and NPV), accuracy, F-measure, likelihood ratio positive (LR+), and likelihood ratio negative (LR-). The true-positive fraction is the fraction of actual rejection properly predicted as rejection (sensitivity). The false-positive fraction is the fraction of actual non-rejection that was inaccurately predicted to rejection (1-specificity). The area under the ROC curve (AUC) was applied as an index of performance. AUCs were calculated on the results obtained from the test set. AUC may range from 0 to 1, with area of 1.0 corresponding to perfect discrimination and area of 0.5 to what is expected by chance alone (27). The optimum cut-off value to predict rejection or non-rejection was selected.

3. Results

3.1. Patients' Characteristics

In a preliminary statistical study, different simple and routine laboratory tests of liver transplant patients including liver function tests, cell blood count, and coagulation tests on days 3, 7, and 14 after transplantation were separately assessed to choose those that were significantly different between acute rejection and non-rejection in liver transplant patients. To make equal comparative analysis among days 3, 7, 14 after transplantation, we selected the following seven statistically significant variables: aspartate aminotransferase (IU/L), alanine aminotransferase (IU/L), alkaline phosphatase (IU/L), total serum bilirubin (mg/dL), direct serum bilirubin (mg/dL), total serum protein (g/dL), and serum calcium (mg/dL). Consequently, these variables were used to initially build the inputs of the artificial neural network and a single output layer representing the probability of acute rejection (Figure 1).

The baseline characteristics of acute rejection and non-rejection in liver transplant recipients are shown in Table 1. The mean (SD) age was 38 (± 17.01) years, and the patients were predominantly men (63.5%). The underlying diagnoses of these patients were hepatitis B cirrhosis (30 patients), cryptogenic cirrhosis (27 patients), primary sclerosing cholangitis (26 patients), autoimmune hepatitis (13 patients), Wilson's disease (10 patients), non-alcoholic steatohepatitis (9 patients), and other (33 patients).

Table 1. Baseline Characteristics of the Acute Rejection and Non-Rejection Liver Transplant Patients[a]

Variables	Day 3 (n = 148)		Day 7 (n = 148)		Day 14 (n = 148)	
	Rejection Patients (n = 34)	Non-Rejection Patients (n = 114)	Rejection Patients (n = 34)	Non-Rejection Patients (n = 114)	Rejection patients (n = 34)	Non-Rejection patients (n = 114)
Demographic						
Age, y	34.7 (13.27)	38.05 (17.98)	34.7 (13.27)	38.05 (17.98)	34.7 (13.27)	38.05 (17.98)
Weight, kg	67 (12.50)	60.62 (18.03)	67 (12.50)	60.62 (18.03)	67 (12.50)	60.62 (18.03)
Gender (male/female)	(26/8)	(72/42)	(26/8)	(72/42)	(26/8)	(72/42)
Laboratory Values						
BUN, mg/dL	26.94 (16.20)	28.56 (14.98)	23.41 (19.67)	19.52 (14.38)	20.97 (13.46)	17.55 (12.49)
Creatinine, mg/dL	1.08 (0.98)	0.98 (0.70)	1.10 (1.19)	1.10 (1.26)	1.17 (1.86)	0.95 (0.52)
Sodium, mEq/L	139.20 (4.12)	137.28 (12.38)	139.26 (6.38)	140.07 (3.51)	137.50 (4.01)	137.70 (5.08)
Potassium, mEq/L	4.22 (0.44)	4.11 (0.40)	4.29 (1.07)	4.03 (0.68)	4.44 (0.61)	4.41 (0.79)
Calcium, mg/dL[b]	8.20 (0.71)	7.69 (0.90)	8.02 (1.15)	7.78 (0.91)	8.57 (0.80)	7.95 (1.57)
ALT, IU/L[b]	976.91 (1426.87)	404.22 (228.71)	377.73 (230.60)	151.35 (97.34)	200.41 (117.12)	87.82 (45.94)
AST, IU/L[b]	795.74 (1337.86)	303.65 (176.11)	134.29 (60.16)	41.73 (10.92)	55.67 (30.67)	35.90 (22.20)
ALP, IU/L[b]	321.94 (210.31)	269.91 (165.14)	392.20 (230.29)	315.67 (168.16)	444.79 (234.45)	341.55 (191.88)
GGT, IU/L	17.84 (4.33)	19.08 (7.16)	15.16 (2.12)	15.47 (3.76)	14.33 (1.75)	14.90 (3.08)
Total bilirubin, mg/dL[b]	4.11 (3.78)	3.45 (2.29)	3.92 (3.82)	2.23 (1.65)	3.93 (4.66)	1.59 (0.99)
Direct bilirubin, mg/dL[b]	2.01 (1.84)	2.52 (2.14)	2.15 (2.25)	1.21 (0.89)	2.30 (3.09)	1.11 (0.84)
Albumin, g/dL	3.01 (0.48)	2.87 (0.41)	3.12 (0.53)	3.01 (0.49)	3.30 (0.48)	3.25 (0.53)
Total protein, g/dL[b]	5.36 (1.07)	4.89 (0.73)	5.32 (1.02)	4.92 (0.83)	5.65 (1.23)	5.57 (1.01)
Hemoglobin, g/dL	9.02 (1.75)	8.62 (1.73)	9.19 (2.03)	9.35 (1.54)	9.87 (1.94)	9.49 (1.87)
Hematocrit, %	27.08 (4.87)	26.06 (4.29)	27.93 (5.95)	27.95 (4.08)	32.14 (9.32)	28.43 (5.11)
INR	1.84 (0.81)	1.76 (0.55)	1.39 (0.37)	1.37 (0.51)	1.24 (0.29)	1.29 (0.50)
Platelet count, 1/μL	88865.7 (90269.6)	67159.6 (71650.8)	87687.1 (78435.1)	60120.17 (43951.7)	152938.2 (94136.0)	145375.4 (93658)
WBC, cells/μL	11468.5 (5935.07)	9988.5 (8938.70)	9755.88 (5617.08)	8334.21 (4840.36)	10126.4 (5211.7)	10723.6 (8768.8)

Abbreviations: ALP, alkaline phosphatase; ALT, Alanine Aminotransferase; AST, aspartate aminotransferase; BUN, Blood urea nitrogen; GGT, γ-glutamyltranspeptidase; INR, international normalized ratio; WBC, white cell.
[a]Values are expressed as mean (SD).
[b]Significant input variables.

In this research, there were 3108 data points (21 variables) and the study population included 148 liver transplant patients with acute rejection and non-rejection, of which 102 were apportioned at random to the training set and the remaining 46 to the testing set. The variables used to build the artificial neural network were compared in the

two sets (Table 2). These markers were not significantly different across the groups (P > 0.05). 23% of the patients (34 persons) were rejection cases, of which 24 episodes were used in the training set and 10 in the testing set.

3.2. Optimal Input Variable Data Sets and Prediction Accuracy

The sensitivity analysis of the seven variables used to initially construct the ANN is illustrated in Figure 2. The analysis indicated that the aspartate aminotransferase (IU/L) and alanine aminotransferase (IU/L) were two variables with the highest influence on the rejection outcome on days 7 and 14. Additionally, alanine aminotransferase (IU/L) and calcium were the most impressive variables in the prediction of acute rejection on day 3 whereas the other variables played minor roles.

Table 3 shows the ANN and LR performance in diagnosing the acute rejection in liver transplant patients compared to the gold standard liver biopsy. Some predictive performance indices such as accuracy, sensitivity, specificity, F-measure, positive and negative predictive values (PPV and NPV), likelihood ratio positive (LR+), likelihood ratio negative (LR-), and AUC on days 3, 7, and 14 regarding ANN and LR use in the training and testing sets at the optimum cut-off point are shown in Table 4. All performance indices were superior in the models developed by ANN to those built by LR on the three mentioned days. AUCs for training data and testing data (Figure 3) were higher in ANN models than in LR models. For training data on days 3, 7, and 14, AUCs were 0.83, 0.98, and 0.93 in ANN models and 0.58, 0.81, and 0.66 in LR models, respectively. For testing data on days 3, 7 and 14, AUCs were 0.61, 0.95, and 0.75 in ANN models and 0.50, 0.79, and 0.65 in LR models, respectively.

Finally, the ANN model on day 7 revealed the best results compared to the models on days 3 and 14, which used seven input variables, recognized correctly eight out of ten acute rejection patients and 34 out of 36 non-rejection ones in the testing set (yielding a mean accuracy, a sensitivity, and a specificity of 95%, 94%, and 94%, respectively). These values were 3 out of 10 and 30 out of 36 in LR test, respectively. Therefore, the results showed that ANN model outperformed the LR model on the three mentioned days.

Furthermore, to find any prognostic interrelationship between MELD values and acute allograft rejection, we assessed MELD values in this study. MELD values had the same effect like TP (the day 7 data) in predicting rejection and non-rejection. Based on pre-transplant MELD values, recipients were categorized as low risk ($\leq$ 15), medium risk (16 - 25), and high risk (> 25). The ANN model along with other selected features recognized correctly eight out of ten acute rejection patients and 34 out of 36 non-rejection ones in the testing set. In the group of recipi-

ents with MELD values less than 15, two rejection and six non-rejection cases were predicted. In the second group with MELD values between 16 and 25, 5 rejection and 17 non-rejection cases were predicted. In the last group with MELD values more than 25, one and 11 rejection and non-rejection cases, respectively, were predicted. Therefore, the recipients with MELD values between 16 and 25 had the highest correlation with rejection.

4. Discussion

Acute Rejection (AR) is the most common form of liver allograft rejection. It usually happens within the first month of transplantation and its diagnosis depends heavily on the use of liver biopsy. Given that liver biopsy provides pivotal information for diagnosis and serves to guide further therapeutic decisions, it is itself a highly risky procedure, particularly in patients with impaired coagulation (9, 28). Thus, it is not appropriate for regular monitoring of liver transplant patients. For these reasons, liver transplant patients would hugely benefit from a reliable, non-invasive method of diagnosing rejection.

Artificial neural networks (ANNs) have been studied in different areas of medical applications, including diagnosis of acute myocardial infarction (29), diagnosis of thyroid function from in vitro laboratory tests (30), prediction of urinary tract infection (17), prediction of influenza vaccination outcome (19), cancer diagnosis (18, 31), prediction of delayed renal allograft function (32), predicting acute graft-vs-host disease after unrelated donor hematopoietic stem cell transplantation in thalassemia patients (33), and predicting cytomegalovirus (CMV) disease after renal transplantation (34). In the liver transplantation field, the use of ANNs has been scrutinized in prediction of fibrosis in hepatitis C virus infected liver transplant recipients (12), prediction of outcomes after liver transplantation (21), and prediction of graft failure (22). Furthermore, ANNs have been designed to predict patient survival after liver transplantation (35) and early transplant failure (36).

Although, there are some previous studies in the field of liver transplantation using ANN, to our knowledge, there is only one published study evaluating ANN in predicting acute rejection following liver transplantation. It should be noted that the variables that they used as input to build the ANN and the ANN model were different from those of our study.

In the present study, we developed an ANN model, based entirely on routine and inexpensive laboratory data, in order to diagnose acute rejection in liver transplant patients. In designing the ANN, in order to make equal comparative analysis among days 3, 7, and 14 after transplanta-

Table 2. Comparison of Clinical Features Between Training Data and Validation Data

	Day 3 (n = 148)			Day 7 (n = 148)			Day 14 (n = 148)		
Variables	Training Set (n = 102)	Testing Set (n = 46)	P Value	Training Set (n = 102)	Testing Set (n = 46)	P Value	Training Set (n = 102)	Testing Set (n = 46)	P Value
ALT, IU/L	532.90 (678.93)	612.31 (964.30)	0.571	208.63 (178.78)	197.30 (193.52)	0.728	114.09 (82.80)	120.02 (94.71)	0.705
AST, IU/L	432.84 (756.69)	538.40 (1115.89)	0.507	63.24 (48.65)	58.28 (40.94)	0.548	40.77 (28.15)	44.84 (30.95)	0.439
ALP, IU/L	292.41 (183.84)	287.95 (170.34)	0.891	322.25 (187.40)	281.54 (129.42)	0.184	391.69 (224.9)	293.0 (157.25)	0.093
Total bilirubin, mg/dL	3.49 (2.59)	3.80 (2.99)	0.533	2.56 (2.24)	2.63 (2.51)	0.867	2.23 (2.62)	2.16 (2.73)	0.892
Direct bilirubin, mg/dL	2.30 (1.95)	2.44 (2.0)	0.691	1.37 (1.26)	1.45 (1.41)	0.716	1.45 (1.74)	1.47 (1.81)	0.953
Total protein, g/dL	4.98 (0.87)	5.09 (0.93)	0.491	5.10 (0.90)	4.90 (0.91)	0.209	5.60 (1.07)	5.52 (1.09)	0.686
Calcium, mg/dL	7.81 (0.81)	7.83 (1.06)	0.897	7.87 (0.90)	7.91 (0.75)	0.800	8.03 (1.55)	8.11 (1.50)	0.767
Rejection/non-rejection	24/78	10/36	0.811	24/78	10/36	0.811	24/78	10/36	0.811

Abbreviations: ALP, alkaline phosphatase; ALT, Alanine Aminotransferase; AST, aspartate aminotransferase.

Figure 2. 'Sensitivity' Analysis of Input Variables on Days 3, 7, and 14; (A) day 3; (B) day 7; (C) day 14
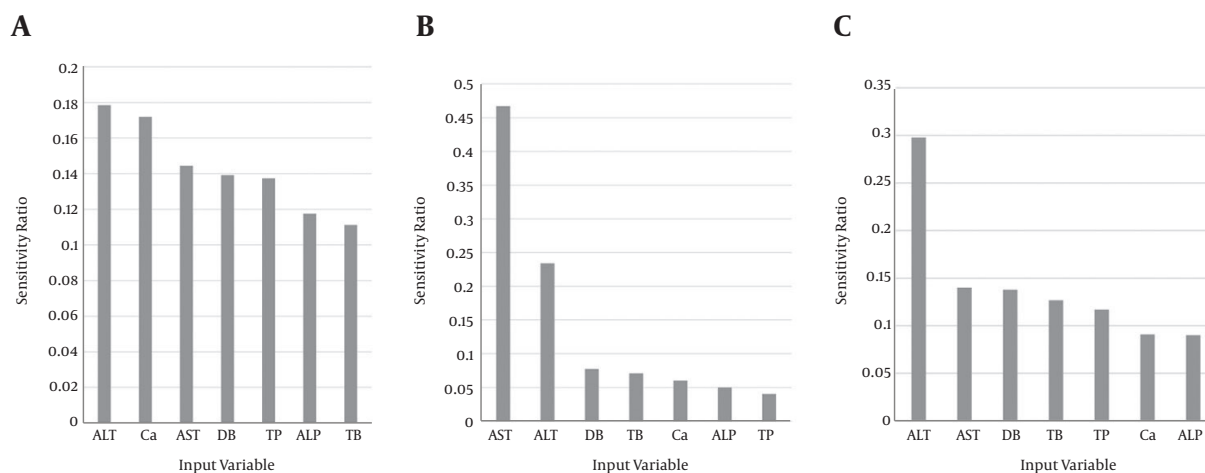
Value shown for each input variable is a measure of its relative importance, with 0 representing a variable that has no effect on the prediction and 1 representing a variable that completely dominates the prediction. AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; TP, total bilirubin; DB, direct bilirubin; TP, total protein; Ca, calcium.

Table 3. The ANN and LR Performance in Diagnosing Acute Rejection and Non-Rejection in Liver Transplant Patients, Compared to Liver Biopsy

		ANN		Logistic Regression	
Liver Biopsy Report		Rejection (N = 10)	Non-Rejection (N = 36)	Rejection (N = 10)	Non-Rejection (N = 36)
Day 3	Positive	3	34	0	36
	Negative	7	2	10	0
Day 7	Positive	8	34	3	35
	Negative	2	2	7	1
Day 14	Positive	5	33	1	36
	Negative	5	3	9	0

tion, seven statistically significant variables were chosen. Moreover, for testing the ANN, we used test data that were not applied for training, and the results were compared with a logistic regression (LR) model using the same data.

The ANN and LR models using optimum cut-off point had the best performance (accuracy 0.90; sensitivity 0.93; specificity 0.90; F-measure 0.94 and accuracy 0.82 sensitivity 0.70, specificity 0.69, F-measure 0.89, respectively) on

Table 4. Comparison of Predictive Performance of ANN (Artificial Neural Network) and Logistic Regression (LR) Using Optimum Cut-Off Point

| | Day 3 (n = 148) | | | | Day 7 (n = 148) | | | | Day 14 (n = 148) | | | |
| | Training Set | | Testing Set | | Training Set | | Testing Set | | Training Set | | Testing Set | |
Measure	ANN	LR	ANN	LR	ANN	LR	ANN	LR	ANN	LR	ANN	LR
Accuracy	0.89	0.76	0.78	0.77	0.95	0.81	0.90	0.82	0.93	0.77	0.83	0.78
Sensitivity	0.72	0.56	0.57	0.51	0.94	0.72	0.87	0.70	0.86	0.59	0.68	0.57
Specificity	0.73	0.55	0.57	0.53	0.94	0.71	0.90	0.69	0.88	0.60	0.71	0.56
PPV, %	0.98	1	0.94	1.0	0.97	0.99	0.95	0.98	0.98	1.0	0.93	1.0
NPV, %	0.55	0	0.23	0.0	0.86	0.23	0.75	0.26	0.78	0.04	0.47	0.02
F-measure	0.93	0.86	0.87	0.87	0.96	0.89	0.94	0.89	0.96	0.87	0.89	0.88
LR+	2.73	1.24	1.13	0.35	0.91	2.56	8.75	0.54	5.38	1.4	1.79	0.40
LR-	0.41	0.82	0.77	0.32	0.05	0.41	0.13	0.17	0.17	0.68	0.45	0.27
AUC	0.83	0.58	0.61	0.50	0.98	0.81	0.95	0.79	0.93	0.66	0.75	0.65
Optimum cut-off point	0.81	0.76	0.84	0.76	0.73	0.76	0.91	0.76	0.81	0.77	0.93	0.77

Abbreviations: LR-, likelihood ratio negative; LR+, likelihood ratio positive; NPV, negative predictive value; PPV, positive predictive value.

day 7, compared to days 3 and 14. Furthermore, the ANN on this day recognized correctly eight out of ten acute rejection patients and 34 out of 36 non-rejection ones in the testing set.

In all training sets and testing sets, accuracy, sensitivity, specificity, F-measure, and AUC were higher in the 3, 7, and 14 days' models built by ANN than in those constructed by LR; this is consistent with other reports in which ANN outperformed LR in both training and testing (37-40).

In this study, ANN demonstrated accuracy in predicting acute rejection in liver transplant recipients at acceptable levels (AUC 0.95), which was related to day 7. Victoria et al. also reported excellent diagnostic accuracies (AUC 0.902 with sensitivity and specificity of 80.0% and 90.1% at the optimum decision threshold, respectively) when training ANN to predict acute rejection in liver transplant recipients over the first 3 months after transplantation using alanine aminotransferase, bilirubin levels, and day posttransplantation (9). Besides, Cucchetti similarly demonstrated superiority of ANN to MELD in predicting mortality of patients with end-stage liver disease (AUC 0.95). Habib et al. showed that the pre-transplant MELD score was inversely correlated with post-transplant survival based on a larger study with a longer follow-up (41). Moon et al. exhibited that the high MELD group had a significantly lower graft and patient survival rate than the low and intermediate MELD groups did (42).

Another study showed that ANN is superior to pediatric and model for end-stage liver disease (PELD-MELD) for predicting poor outcome in pediatric acute liver failure (PALF) with AUC 0.96, sensitivity 82.6%, and specificity 96% (41). Ibanez et al. also designed an LR and an ANN model to predict early transplant failure using variables from donor, recipient, and operative data. The AUC of this model in the training and validation groups was 75% and 68.7 % for lo-

gistic regression and 96% and 69.9 % for neural network, respectively (36).

The ANN assessed in this study was not developed to replace clinical judgment or liver biopsy, but rather to serve as a help in decision-making in a similar manner to other liver function test results for the diagnosis of acute rejection in the early postoperative period. It can be seen in Table 3 that the ANN output pattern is very similar to that of liver biopsy report after transplantation. Thus, it would be easy for physicians to realize and apply the results obtained from this network.

The two limitations of our study were the small sample size and the short follow-up period of our patients. These results are promising, particularly when compared to the performance of traditional statistical models on the same data set. The freedom of neural networks from a priori assumptions concerning the data, and their robustness when confronting with noisy data generated by nonlinear processes make them encouraging and well-suited tools through which we can develop predictive clinical models.

In conclusion, this research suggests that the ANN may be as a potential clinical decision-support tool for identifying liver transplant recipients at risk of rejection using clinical information and simple laboratory tests. In future research, Cohort studies using higher sample sizes, and long follow-up periods of patients could improve such new models in decision-making and patient management in actual clinical practice.

In conclusion, this research showed the proposed predictive models make an acceptable predictive power and they can be used as support methods for facilitating clinical decision-making processes and disease prognosis. In future studies, cohort studies using higher sample sizes and long follow-up periods of patients could improve such models in decision-making and patient management in ac-

Figure 3. ROC Curves and AUCs ($\pm$ SD) for ANN and LR Models of Days 3, 7, and 14 in the Two Sets
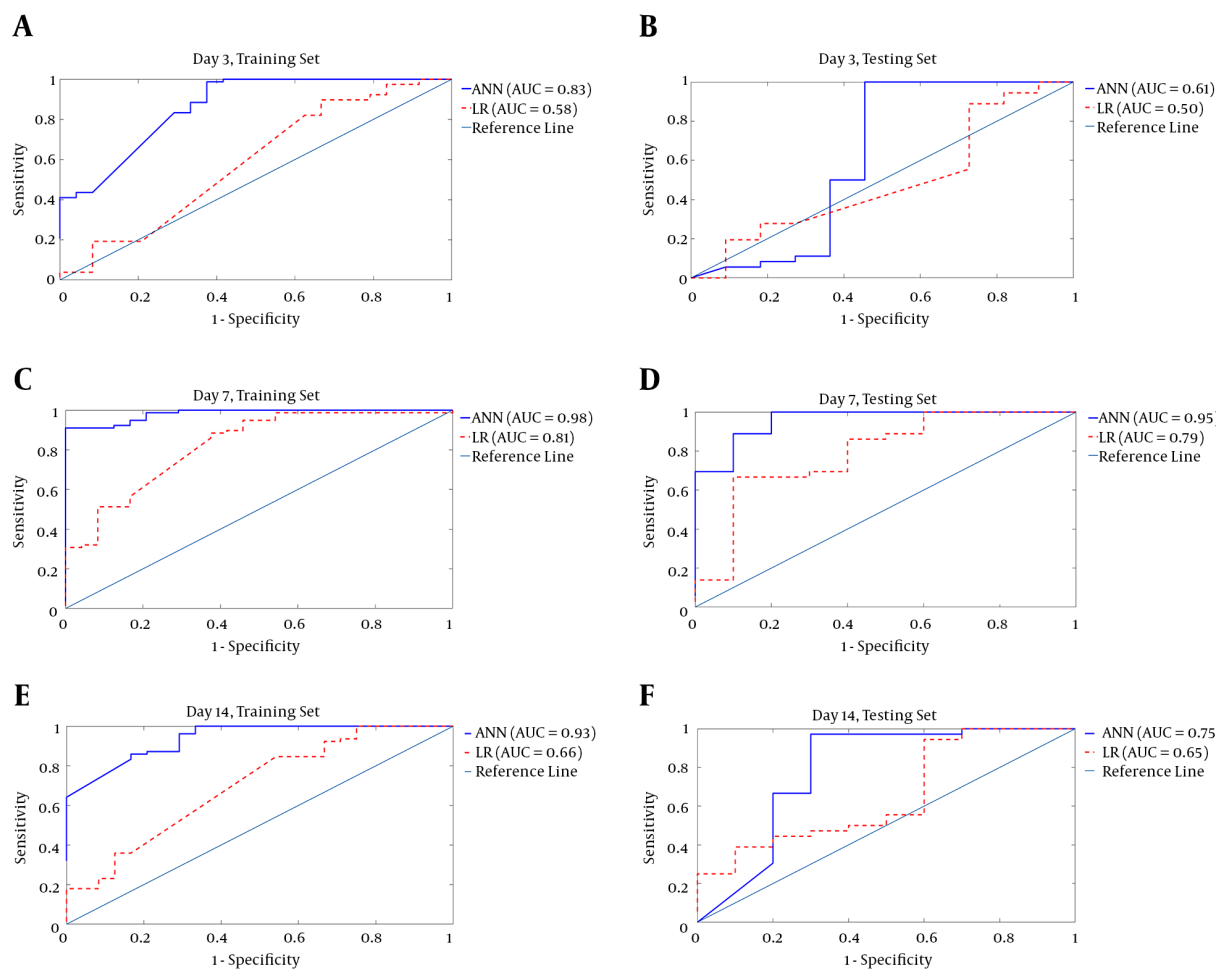
A, in the training set AUCs for day 3 (ANN = 0.83 $\pm$ 0.12, LR = 0.58 $\pm$ 0.03); B, in the testing set AUCs for day 3 (ANN = 0.61 $\pm$ 0.09, LR = 0.50 $\pm$ 0.10); C, In the training set AUCs for day 7 (ANN = 0.98 0.01, LR = 0.81 $\pm$ 0.01); D, In the testing set AUCs for day 7 (ANN = 0.95 $\pm$ 0.03, LR = 0.79 $\pm$ 0.04); E, In the training set AUCs for day 14 (ANN = 0.93 0.07, LR = 0.66 $\pm$ 0.02); F, In the testing set AUCs for day 14 (ANN = 0.75 $\pm$ 0.10 , LR = 0.65 $\pm$ 0.07).

tual clinical practice.

Footnotes

Authors' Contribution: Study concept, design, and drafting of the manuscript: Abdolhossein Zare and Negar Azarpira; acquisition of data: Abdolhossein Zare and Neda Zarei; analysis and interpretation of data: Abdolhossein Zare, Mohammad Amin Zare, Ramin Yaghoobi, Saeede Salehi and Mohammad Ali Zare; critical revision: Bita Geramizadeh, Seid Ali Malekhosseini and Negar Azarpira.

References

1. Zare A, Karimi MH, Rashki A, Geramizadeh B, Afshari A, Miri HR, et al. Association of the Interleukin-27 Gene Expression and Hepatitis B Virus Infection in Liver Transplanted Patients. *Exp Clin Transplant*. 2017;**15**(5):554–60. doi: 10.6002/ect.2015.0243. [PubMed: 26925776].

2. Marzano A. Management of HBV resistance in the post-transplant setting. *Digest Liver Dis Suppl*. 2011;**5**(1):19–22. doi: 10.1016/s1594-5804(11)60020-x.

3. Wu CF, Wu YJ, Liang PC, Wu CH, Peng SF, Chiu HW. Disease-free survival assessment by artificial neural networks for hepatocellular carcinoma patients after radiofrequency ablation. *J Formos Med Assoc*. 2017;**116**(10):765–73. doi: 10.1016/j.jfma.2016.12.006. [PubMed: 28117199].

4. Khosravi B, Pourahmad S, Bahreini A, Nikeghbalian S, Mehrdad G. Five Years Survival of Patients After Liver Transplantation and Its Effective Factors by Neural Network and Cox Poroportional Hazard Regression Models. *Hepat Mon*. 2015;**15**(9). e25164. doi: 10.5812/hepatmon.25164. [PubMed: 26500682].

5. Wiesner RH, Menon KV. Late hepatic allograft dysfunction. *Liver Transpl*. 2001;**7**(11 Suppl 1):S60–73. doi: 10.1053/jlts.2001.29094. [PubMed: 11689778].

6. Rodriguez-Peralvarez M, Rico-Juri JM, Tsochatzis E, Burra P, De la Mata M, Lerut J. Biopsy-proven acute cellular rejection as an efficacy endpoint of randomized trials in liver transplantation: a systematic review and critical appraisal. *Transpl Int*. 2016;**29**(9):961–73. doi: 10.1111/tri.12737. [PubMed: 26714264].

7. Bolognesi M, Sacerdoti D, Mescoli C, Nava V, Bombonato G, Merkel C, et al. Acute liver rejection: accuracy and predictive values of doppler US measurements–initial experience. *Radiology*. 2005;**235**(2):651–8. doi: 10.1148/radiol.2352040506. [PubMed: 15770040].

8. Batts KP. Acute and chronic hepatic allograft rejection: pathology and classification. *Liver Transplant*. 1999;**5**(4 Suppl 1):S21–9.

9. Hughes V. Clinical validation of an artificial neural network trained to identify acute allograft rejection in liver transplant recipients. *Liver Transplant*. 2001;**7**(6):496–503. doi: 10.1053/jlts.2001.24642.

10. Wang D, Wang Q, Shan F, Liu B, Lu C. Identification of the risk for liver fibrosis on CHB patients using an artificial neural network based on routine and serum markers. *BMC Infect Dis*. 2010;**10**:251. doi: 10.1186/1471-2334-10-251. [PubMed: 20735842].

11. Raoufy MR, Vahdani P, Alavian SM, Fekri S, Eftekhari P, Gharibzadeh S. A novel method for diagnosing cirrhosis in patients with chronic hepatitis B: artificial neural network approach. *J Med Syst*. 2011;**35**(1):121–6. doi: 10.1007/s10916-009-9348-8. [PubMed: 20703578].

12. Piscaglia F, Cucchetti A, Benlloch S, Vivarelli M, Berenguer J, Bolondi L, et al. Prediction of significant fibrosis in hepatitis C virus infected liver transplant recipients by artificial neural network analysis of clinical factors. *Eur J Gastroenterol Hepatol*. 2006;**18**(12):1255–61. doi: 10.1097/01.meg.0000243885.55562.7e. [PubMed: 17099373].

13. Shaikhina T, Khovanova NA. Handling limited datasets with neural networks in medical applications: A small-data approach. *Artif Intell Med*. 2017;**75**:51–63. doi: 10.1016/j.artmed.2016.12.003. [PubMed: 28363456].

14. Desai KM, Vaidya BK, Singhal RS, Bhagwat SS. Use of an artificial neural network in modeling yeast biomass and yield of β-glucan. *Proc Biochem*. 2005;**40**(5):1617–26. doi: 10.1016/j.procbio.2004.06.015.

15. Cucchetti A, Vivarelli M, Heaton ND, Phillips S, Piscaglia F, Bolondi L, et al. Artificial neural network is superior to MELD in predicting mortality of patients with end-stage liver disease. *Gut*. 2007;**56**(2):253–8. doi: 10.1136/gut.2005.084434. [PubMed: 16809421].

16. Dorado-Moreno M, Perez-Ortiz M, Gutierrez PA, Ciria R, Briceno J, Hervas-Martinez C. Dynamically weighted evolutionary ordinal neural network for solving an imbalanced liver transplantation problem. *Artif Intell Med*. 2017;**77**:1–11. doi: 10.1016/j.artmed.2017.02.004. [PubMed: 28545607].

17. Heckerling PS, Canaris GJ, Flach SD, Tape TG, Wigton RS, Gerber BS. Predictors of urinary tract infection based on artificial neural networks and genetic algorithms. *Int J Med Inform*. 2007;**76**(4):289–96. doi: 10.1016/j.ijmedinf.2006.01.005. [PubMed: 16469531].

18. Bottaci L, Drew PJ, Hartley JE, Hadfield MB, Farouk R, Lee PW, et al. Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. *Lancet*. 1997;**350**(9076):469–72. doi: 10.1016/S0140-6736(96)11196-X. [PubMed: 9274582].

19. Trtica-Majnaric L, Zekic-Susac M, Sarlija N, Vitale B. Prediction of influenza vaccination outcome by neural networks and logistic regression. *J Biomed Inform*. 2010;**43**(5):774–81. doi: 10.1016/j.jbi.2010.04.011. [PubMed: 20451660].

20. Myers J, de Souza CR, Borghi-Silva A, Guazzi M, Chase P, Bensimhon D, et al. A neural network approach to predicting outcomes in heart failure using cardiopulmonary exercise testing. *Int J Cardiol*. 2014;**171**(2):265–9. doi: 10.1016/j.ijcard.2013.12.031. [PubMed: 24387896].

21. Doyle HR, Dvorchik I, Mitchell S, Marino IR, Ebert FH, McMichael J, et al. Predicting outcomes after liver transplantation. A connectionist approach. *Ann Surg*. 1994;**219**(4):408–15. [PubMed: 8161267].

22. Matis S, Doyle H, Marino I, Mural R, Uberbacher E. Use of neural networks for prediction of graft failure following liver transplantation. *Proceedings of the Eighth IEEE Symposium on; IEEE*. 1995. p. 133–40.

23. Wight DGD, Portmann B. Pathology of liver transplantation. *System Pathol*. 1987;**11**:543–96.

24. Raji CG, Vinod Chandra SS. Graft survival prediction in liver transplantation using artificial neural network models. *J Comput Sci*. 2016;**16**:72–8. doi: 10.1016/j.jocs.2016.05.005.

25. Azar AT. Fast neural network learning algorithms for medical applications. *Neur Comput Appl*. 2012;**23**(3-4):1019–34. doi: 10.1007/s00521-012-1026-y.

26. Hanley JA. Receiver operating characteristic (ROC) methodology: the state of the art. *Crit Rev Diagn Imaging*. 1989;**29**(3):307–35. [PubMed: 2667567].

27. Banerjee R, Das A, Ghoshal UC, Sinha M. Predicting mortality in patients with cirrhosis of liver with application of neural network technology. *J Gastroenterol Hepatol*. 2003;**18**(9):1054–60. [PubMed: 12911662].

28. Jiang W, Farber JL. *Pathology of Liver Transplantation 21. Contemporary Liver Transplantation*. 2017.

29. Pedersen SM, Jorgensen JS, Pedersen JB. Use of neural networks to diagnose acute myocardial infarction. II. A clinical application. *Clin Chem*. 1996;**42**(4):613–7. [PubMed: 8605680].

30. Sharpe PK, Solberg HE, Rootwelt K, Yearworth M. Artificial neural networks in diagnosis of thyroid function from in vitro laboratory tests. *Clin Chem*. 1993;**39**(11 Pt 1):2248–53. [PubMed: 8222217].

31. Kuo RJ, Huang MH, Cheng WC, Lin CC, Wu YH. Application of a two-stage fuzzy neural network to a prostate cancer prognosis system. *Artif Intell Med*. 2015;**63**(2):119–33. doi: 10.1016/j.artmed.2014.12.008. [PubMed: 25576196].

32. Brier ME, Ray PC, Klein JB. Prediction of delayed renal allograft function using an artificial neural network. *Nephrol Dial Transplant*. 2003;**18**(12):2655–9. [PubMed: 14605292].

33. Caocci G, Baccoli R, Vacca A, Mastronuzzi A, Bertaina A, Piras E, et al. Comparison between an artificial neural network and logistic regression in predicting acute graft-vs-host disease after unrelated donor hematopoietic stem cell transplantation in thalassemia patients. *Exp Hematol*. 2010;**38**(5):426–33. doi: 10.1016/j.exphem.2010.02.012. [PubMed: 20206661].

34. Sheppard D, McPhee D, Darke C, Shrethra B, Moore R, Jurewitz A, et al. Predicting cytomegalovirus disease after renal transplantation: an artificial neural network approach. *Int J Med Inform*. 1999;**54**(1):55–76. [PubMed: 10206429].

35. Cruz-Ramirez M, Hervas-Martinez C, Fernandez JC, Briceno J, de la Mata M. Predicting patient survival after liver transplantation using evolutionary multi-objective artificial neural networks. *Artif Intell Med*. 2013;**58**(1):37–49. doi: 10.1016/j.artmed.2013.02.004. [PubMed: 23489761].

36. Ibanez V, Pareja E, Serrano AJ, Vila JJ, Perez S, Martin JD, et al. Predicting Early Transplant Failure: Neural Network Versus Logistic Regression Models. *Open Transplant J*. 2009;**3**(1):14–21. doi: 10.2174/1874418400903010014.

37. Tangri N, Ansell D, Naimark D. Predicting technique survival in peritoneal dialysis patients: comparing artificial neural networks and logistic regression. *Nephrol Dial Transplant*. 2008;**23**(9):2972–81. doi: 10.1093/ndt/gfn187. [PubMed: 18441002].

38. Green M, Bjork J, Forberg J, Ekelund U, Edenbrandt L, Ohlsson M. Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room. *Artif Intell Med*. 2006;**38**(3):305–18. doi: 10.1016/j.artmed.2006.07.006. [PubMed: 16962295].

39. Eftekhar 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.

[PubMed: 15713231].

40. Chiu HC, Ho TW, Lee KT, Chen HY, Ho WH. Mortality predicted accuracy for hepatocellular carcinoma patients with hepatic resection using artificial neural network. *ScientificWorldJournal*. 2013;**2013**:201976. doi: 10.1155/2013/201976. [PubMed: 23737707].

41. Moon DB, Lee SG, Kang WH, Song GW, Jung DH, Park GC, et al. Adult Living Donor Liver Transplantation for Acute-on-Chronic Liver Failure in High-Model for End-Stage Liver Disease Score Patients. *Am J Transplant*. 2017;**17**(7):1833–42. doi: 10.1111/ajt.14198. [PubMed: 28097804].

42. Rajanayagam J, Frank E, Shepherd RW, Lewindon PJ. Artificial neural network is highly predictive of outcome in paediatric acute liver failure. *Pediatr Transplant*. 2013;**17**(6):535–42. doi: 10.1111/petr.12100. [PubMed: 23802584].