JOURNAL OF VIRAL HEPATITIS
JVH

Journal of Viral Hepatitis, 2013, **20,** 248–255 doi:10.1111/j.1365-2893.2012.01647.x

A model to predict 3-month mortality risk of acute-on-chronic hepatitis B liver failure using artificial neural network

M.-H. Zheng,[1,2,*] K.-Q. Shi,[1,2,*] X.-F. Lin,[1,3,*] D.-D. Xiao,[1,3] L.-L. Chen,[1,2] W.-Y. Liu,[1,3] Y.-C. Fan[4] and Y.-P. Chen[1,2] [1]*Department of Infection and Liver Diseases, Liver Research Center, The First Affiliated Hospital of Wenzhou Medical College, Wenzhou, China;* [2]*Institute of Hepatology, Wenzhou Medical College, Wenzhou, China;* [3]*School of the First Clinical Medical Sciences, Wenzhou Medical College, Wenzhou, China; and* [4]*Department of Hepatology, Qilu Hospital of Shandong University, Jinan, China*

Received May 2012; accepted for publication June 2012

SUMMARY. Model for end-stage liver disease (MELD) scoring was initiated using traditional statistical technique by assuming a linear relationship between clinical features, but most phenomena in a clinical situation are not linearly related. The aim of this study was to predict 3-month mortality risk of acute-on-chronic hepatitis B liver failure (ACHBLF) on an individual patient level using an artificial neural network (ANN) system. The ANN model was built using data from 402 consecutive patients with ACHBLF. It was trained to predict 3-month mortality by the data of 280 patients and validated by the remaining 122 patients. The area under the curve of receiver operating characteristic (AUROC) was calculated for ANN and MELD-based scoring systems. The following variables age ($P < 0.001$), prothrombin activity ($P < 0.001$), serum sodium ($P < 0.001$), total bilirubin ($P = 0.015$), hepatitis B e antigen positivity rate ($P < 0.001$) and haemoglobin ($P < 0.001$) were significantly related to the prognosis of ACHBLF and were selected to build the ANN. The ANN performed significantly better than MELD-based scoring systems both in the training cohort (AUROC = 0.869 *vs* 0.667, 0.591, 0.643, 0.571 and 0.577; $P < 0.001$, respectively) and in the validation cohort (AUROC = 0.765 *vs* 0.599, 0.563, 0.601, 0.521 and 0.540; $P \leq 0.006$, respectively). Thus, the ANN model was shown to be more accurate in predicting 3-month mortality of ACHBLF than MELD-based scoring systems.

Keywords: artificial neural network, hepatitis B virus, live disease, prognostic model.

INTRODUCTION

Hepatitis B virus (HBV) is a major human pathogen that causes high morbidity and mortality worldwide. HBV is one of the leading causes for rapid deterioration of liver function, which is a severe condition termed as 'acute-on-chronic liver failure (ACLF)' with high mortality [1]. There is a high prevalence of HBV in the developing countries in Asian where acute-on-chronic hepatitis B liver failure (ACHBLF) accounts for more than 70% of ACLF and almost 120 000 patients die of ACHBLF each year [1,2]. Liver transplantation is the only effective therapeutic option for ACHBLF patients. However, liver transplantation cannot be extensively applied because of the shortage of liver donors and other socioeconomic problems [3]. Thus, an early predictive model, which is objective, reasonable and accurate, is necessary for severity discrimination and organ allocation to decrease the mortality of ACHBLF.

So far, model of end-stage liver disease (MELD) scoring has been widely used for the prediction of live diseases. MELD is initially applied to predict the mortality risk of patients with end-stage liver disease undergoing transjugular intrahepatic portosystemic stent shunt [4]. Based on MELD score, serum sodium and model for MELD (MELD-Na), MELD to serum sodium ratio (MESO), incorporating serum sodium and age MELD model (iMELD), etc., have been gradually proposed [5–7] to predict the prognosis of

Abbreviations: ACHBLF, acute-on-chronic hepatitis B liver failure; ACLF, acute-on-chronic liver failure; AFP, alpha-fetoprotein; ANN, artificial neural network; AUROC, The area under the curve of receiver operating characteristic; BP, back propagation; INR, international normalized ratio; MELD, model of end-stage liver disease; MESO, MELD to serum sodium ratio; MLP, multilayer perceptron; PTA, prothrombin activity.

Correspondence: Ming-Hua Zheng and Yong-Ping Chen, Department of Infection and Liver Diseases, Liver Research Center, the First Affiliated Hospital of Wenzhou Medical College; Institute of Hepatology, Wenzhou Medical College, No. 2 Fuxue lane, Wenzhou 325000, China. E-mails: blueman1320@163.com; ypchen106@yahoo.com.cn
*These authors contributed equally to this paper.

patients with various severe liver diseases in recent years [8–10]. Nevertheless, MELD-based scoring systems still failed to predict the mortality of a considerable proportion of patients, and their predictive accuracy was not satisfying enough [11–13].

All of the above models are constructed on the basis of multiple linear regressions. However, the human body is a complex biological system, and most of the clinical features have multidimensional and nonlinear relationship. So it is ideally difficult to predict the prognosis of liver diseases with a conventional statistical technique. The artificial neural network (ANN) is a novel computer model inspired by the working of the human brain. It can build nonlinear statistical models to deal with complex biological systems. In recent years, ANN models have been introduced in clinical medicine for clinical validations, including predicting hepatocellular carcinoma patients' disease-free survival and preoperative tumour grade [14,15], predicting the mortality of patients with end-stage liver disease [16] and identifying the risk of prostate carcinoma [17]. The ANN model has been proved more accurate and performs better than multiple logistic regression or multiple linear discriminant analysis models [17,18].

Compared with decompensated cirrhosis, ACHBLF is a relatively independent system [1]. As current prognostic models applied in ACHBLF are still not ideal, a suitable and special predictive model is urgently needed. The aim of the present study was to establish an ANN model to predict 3-month mortality of ACHBLF and test its validation.

PATIENTS AND METHODS

Patients

From January 2007 to September 2010, 583 consecutive patients (416 men; 167 women; mean age, 48.2 ± 12.6 years) were diagnosed with ACLF at the First Affiliated Hospital of Wenzhou Medical College. Patients with ACHBLF were included in the study. ACHBLF was defined as an acute hepatic insult manifesting as jaundice and coagulopathy, complicated within 4 weeks by ascites and/or encephalopathy in a patient with chronic HBV infection according to consensus recommendations of the Asian Pacific Association for the Study of the Liver in 2009 [1]. Patients with evidence of non-B hepatitis virus, alcohol abuse, autoimmune, toxic or other causes that might lead to liver failure, past or current hepatocellular carcinoma, liver transplantation or serious diseases in other organ systems were excluded (Fig. 1).

Informed consent was obtained from each patient included in the study, and the research protocol was approved by the Ethics Committee of the First Affiliated Hospital of Wenzhou Medical College. The study procedure conformed with the Helsinki Declaration and Standards for the Reporting of Diagnostic Accuracy Studies [19].

Follow-up

The routine therapy of these patients was the same, including absolute bed rest, energy supplements and vitamins, intravenous infusion of albumin, maintenance water, electrolyte and acid-base equilibrium, prevention and treatment complications [1]. The start date of the follow-up was the date of the diagnosis of ACHBLF. In this study, patients receiving liver transplantation within 3 months were considered as dead. All patients with ACHBLF were followed up for at least 3 months, and the outcome (death or survival) of corresponding patient was recorded.

Laboratory tests

Liver function tests, complete blood count and coagulation tests were performed within the first 24 h after admission. The liver function tests included alanine aminotranferase, aspartate aminotranferase, total bilirubin (TBil), albumin, serum sodium, alpha-fetoprotein (AFP) and creatinine. Complete blood count was made up of platelet and haemoglobin (Hb). Coagulation tests contained prothrombin activity (PTA) and international normalized ratio (INR). Additionally, hepatitis B e antigen (HBeAg) was detected by conventional serological assays. Serum HBV DNA was measured by quantitative PCR assay (Roche Amplicor, limit of detestability of 100 IU/mL) after admission.

MELD-based Scoring Systems

MELD score ($R = 9.57 \times \ln(\text{creatinine (mg/dL)}) + 3.78 \times \ln(\text{bilirubin (mg/dL)}) + 11.2 \times \ln(\text{INR}) + 6.43$) was used to measure the mortality risk in patients with end-stage liver disease [4]. Given the lack of donors, MELD was used as organ allocation tool to increase graft success rate and patient survival rates, which was generally accepted [20]. Recently, some adjustments were added to the original MELD formula to overcome limitations of MELD score. Published data suggested that MELD-Na ($R = \text{MELD} + 1.59 \times (135 - \text{serum sodium (mM)})$) might improve the prognostic accuracy [5]. Furthermore, several other scoring systems such as MELDNa ($R = \text{MELD} - \text{serum sodium (mM)} - (0.025 \times \text{MELD} \times (140 - \text{serum sodium (mM)})) + 140$), MESO ($R = (\text{MELD/serum sodium (mM)}) \times 100$), iMELD ($R = \text{MELD} + (\text{age (year)} \times 0.3) - (0.7 \times \text{serum sodium (mM)}) + 100$)) had been described for predicting the mortality of end-stage liver disease accurately [6,7,21].

Construction of ANN

An ANN can mimic a biological neural system both structurally and functionally. It consists of a set of highly complex, interconnected processing units (neurones) linked with weighted connections and include an input layer, an output layer and one or more hidden layers [22–25]. The input

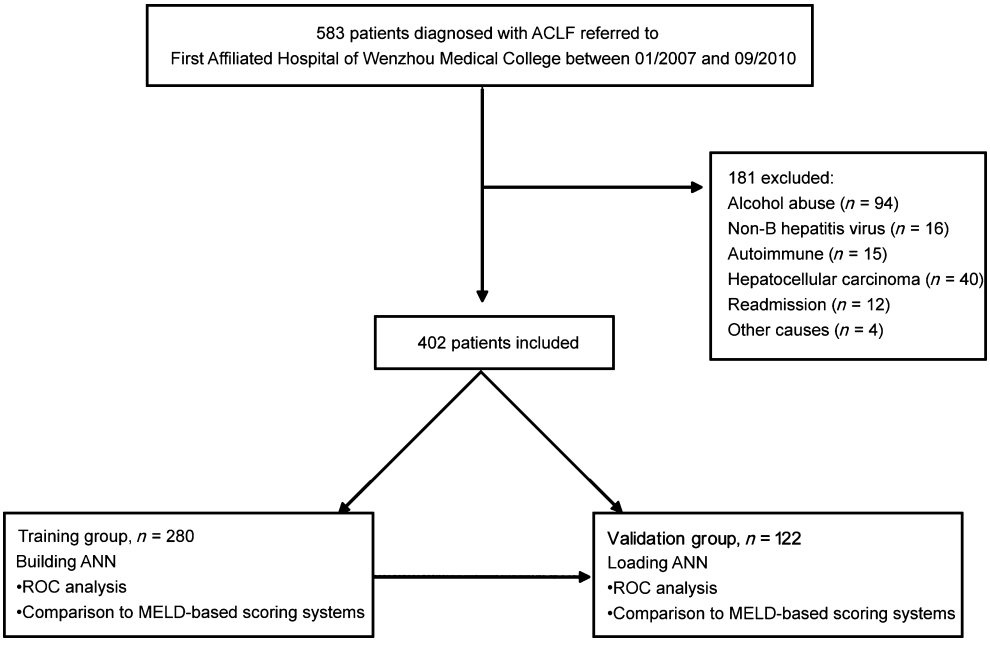

Fig. 1 Data sets used in the training and validating the artificial neural network.

layer contains neurones that receive the data available for the analysis (e.g. various clinical, demographic or laboratory data), and the output layer contains neurones that export different predictive outcomes (e.g. clinical diagnosis or prognosis). The hidden layers are used to allow complex relations between the input and output neurones to evolve.

The advantages of ANN over conventional statistical techniques include three aspects, that is, self-learning process, self-adapting process and inference process. First, it can perform self-learning process. ANN can learn through examples and associate each input with the corresponding output by modifying the weight of the connections between neurones. Once an input is applied as a stimulus to the first layer of neurones, it will be propagated through each upper layer until an output is produced. Afterwards, a self-adapting process is carried out. The output value is compared with the desired output. If there is a discrepancy between these two values, an error signal is generated and then a back propagation (BP) method is applied to alter the weight of the connections between neurones to decrease the overall error of the network. As learning proceeds, the error between the ANN output and the desired output decreases until a minimum is reached (convergence of the network). After these two training processes, the ANN can generate outputs (prognosis) from new input data based on the knowledge accumulated during training, which is regarded as inference process. Thus, after training, the ANN can identify patterns or make predictions on data sets never seen before [22–25].

In this study, we built an ANN by using the graphical neural network development tool NeuroSolution V5.05 (Neurodimension, Gainesville, FL, USA). Variables found to be significantly related to the prognosis of patients with ACHBLF were selected to build the ANN. Four hundred and two eligible patients were assigned to a training cohort (n = 280; 70%) and a validation cohort (n = 122; 30%) randomly. One of the major limitations of ANN is over-training, which can lead to good performance on training sets but poor performance on relatively independent validation sets. To avoid over-training during building of the ANN, 196 patients (70%) were again randomly selected from the training group to train the network and the remaining 84 (30%) were used for cross-validation.

The learning mechanism applied on this ANN was BP by calculating the errors between output value and desired output value. Then, the weight of the connections was altered between neurones to decrease the overall errors of the network. Training was terminated when the sum of square errors was at minimum, compared with the cross-validation data set. The activation function, representing the outcomes of ANN, was used with continuous outputs on the interval from 0 to 1, in which 0 = survival, 1 = death.

Statistical analysis

Statistical analysis was performed using SPSS 13.0 software (SPSS Inc, Chicago, IL, USA) and MedCalc 10.0 software (Mariakerke, Belgium). The Kolmogorov–Smirnov test was applied to determine whether sample data were likely to be derived from a normal distribution population. Continuous variables were expressed by mean ± standard deviation and compared using Wilcoxon signed rank test or Mann–Whitney *U*-test when necessary. Categorical variables were described by proportions or count and compared using proportions chi-square test or the Fisher's exact test when necessary.

Univariate analysis was applied to assess the relationship between clinical or biochemical parameters (input variables) and prognosis (output variables). The variables we selected as the input layer to build the ANN for predicting the prognosis of patients were required to meet the following criteria: important clinical features and statistically significant difference. Besides, odds ratio (OR) and corresponding 95% confidence interval (CI) were calculated.

Performances of the ANN prediction in the training cohort and in the validation cohort were tested using receiver operating characteristic curve (ROC) analysis, in which area under the ROC curve (AUROC) was used to compare the performance of the ANN and MELD-based scoring series using the Hanley and McNeil method. A value of $P < 0.05$ was considered statistically significant.

RESULTS

Baseline characteristic of patients

583 patients diagnosed as ACLF were enrolled in the study. Of these, 181 patients were excluded (Fig. 1). Finally, a total of 402 patients were included. The population characteristics, biochemical characteristics and MELD-based scoring systems of the 402 patients are presented in Table 1. In all the patients, the mean age was 47.2 ± 13.3 years, and the male proportion was 78.6%. One hundred and sixty patients died over a 3-month follow-up (39.8%). The mean MELD, MELD-Na, MELDNa, MESO and iMELD scores of the entire study population were 28.2 ± 6.2, 38.8 ± 14.1, 30.4 ± 5.7, 29.5 ± 21.2, 59.7 ± 31.8, respectively. Patients who survived had lower MELD scores (27.1 ± 6.0 *vs* 29.8 ± 6.2, $P < 0.001$), TBil (16.4 ± 9.8 mg/dL *vs* 18.9 ± 9.8 mg/dL, $P = 0.014$), INR value (2.7 ± 1.9 *vs* 3.1 ± 1.3, $P = 0.035$) and higher serum sodium (135.4 ± 5.1 mM *vs* 131.8 ± 6.5 mM, $P < 0.001$), AFP (207.3 ± 358.2 μg/L *vs* 142.4 ± 256.5 μg/L, $P = 0.035$), Hb (120.7 ± 21.7 g/L *vs* 112.2 ± 24.3 g/L, $P < 0.001$), PTA (25.7 ± 15.3% *vs* 15.5 ± 15.2%, $P < 0.001$) and HBeAg positivity rate (63.2% *vs* 45.0%, $P = 0.019$) (Table S1). Specific characteristics of the training and validation cohorts used to build and test the ANN are described in Table 2 and Table S2.

Construction of ANN

As shown in Table 3, serum sodium (OR = 0.899, 95% CI: 0.866–0.932, $P < 0.001$), TBil (OR = 1.026, 95% CI: 1.005–1.047, $P = 0.015$), age (OR = 1.039, 95% CI: 1.022–1.055, $P < 0.001$), PTA (OR = 0.955, 95% CI: 0.941–0.970, $P < 0.001$), Hb (OR = 0.984, 95% CI: 0.975–0.993, $P < 0.001$) and HBeAg positivity rate (OR = 2.101, 95% CI: 1.399–3.155, $P < 0.001$) were significantly associated with 3-month mortality at the univariate analysis which were all used to build the ANN.

Table 1 Baseline characteristics of the study population

Variables	All patients (n = 402)
Death (%)	160 (39.8)
Age (years)	47.2 ± 13.3
Male gender (%)	316 (78.6)
ALT (IU/L)	204.9 ± 489.1
AST (IU/L)	176.1 ± 310.1
AST/ALT	1.2 ± 1.1
Serum sodium (mM)	134.0 ± 6.0
TBil (mg/dL)	17.4 ± 9.9
Albumin (g/L)	29.5 ± 5.5
Creatinine (mg/dL)	0.9 ± 0.6
Platelet (g/L)	101.9 ± 56.8
Hb (g/L)	117.3 ± 13.3
AFP (μg/L)	181.5 ± 322.8
PTA (%)	21.7 ± 16.0
INR	2.9 ± 1.7
HBV DNA (log10 IU/mL)	5.6 ± 1.6
HBeAg (positivity rate, %)	225 (56.0)
MELD-based scoring systems	
MELD	28.2 ± 6.2
MELD-Na	38.8 ± 14.1
MELDNa	30.4 ± 5.7
MESO	29.5 ± 21.2
iMELD	59.7 ± 31.8

AFP, alpha-fetoprotein; ALT, alanine aminotranferase; AST, aspartate aminotranferase; Hb, hemoglobin; HBeAg, hepatitis B e antigen; iMELD, incorporating serum sodium and age MELD model; INR, international normalized ratio; MELD, the model for end-stage liver disease; MELD-Na, serum sodium and model for MELD; MESO, MELD to serum sodium ratio; PTA, prothrombin activity; TBil, total bilirubin.

Multilayer perceptron (MLP) is one of the most popular and mature ANN architectures with a feed forward neural network where processing neurones are grouped into layers and connected by weighted links. We therefore established an ANN model using MLP. In this present study, MLP included the input, hidden and output layers. Neurones were linked with weighted connections (Fig. 2). In general, the number of input variables and output variables were respectively equal to the number of clinical or biochemical parameters and prognosis we set. As Fig. 2 shows, the MLP has six input neurones and one output neurone. After the debugging and testing for many times, three hidden neurones were added to the hidden layer to increase the MLP's performance.

Assessment of the predictive accuracy of ANN compared with MELD-based scoring systems

In the training cohort, the accuracy of the ANN in predicting 3-month mortality was high (AUROC = 0.869, 95%

Table 2 Baseline characteristics of the study population stratified by ANN cohorts

Variables	Training cohort $n = 280$	Validation cohort $n = 122$	P
Death (%)	114 (40.7)	46 (37.7)	0.571
Age (years)	47.2 ± 13.0	47.4 ± 14.0	0.854
Male gender (%)	219 (78.2)	97 (79.5)	0.771
ALT (IU/L)	198.6 ± 493.3	219.5 ± 481.1	0.694
AST (IU/L)	168.7 ± 306.3	193.2 ± 319.3	0.467
AST/ALT	1.25 ± 1.2	1.1 ± 0.6	0.177
Serum sodium (mM)	133.8 ± 6.0	134.2 ± 5.8	0.577
TBil (mg/dL)	17.6 ± 9.9	17.1 ± 9.8	0.639
Albumin (g/L)	29.6 ± 5.6	29.4 ± 5.1	0.765
Creatinine (mg/dL)	0.9 ± 0.5	0.9 ± 0.748	0.416
Platelet (g/L)	99.3 ± 54.7	107.8 ± 61.2	0.171
Hb (g/L)	118.0 ± 23.0	115.8 ± 23.2	0.390
AFP (μg/L)	167.7 ± 292.2	213.1 ± 383.4	0.244
PTA (%)	20.6 ± 15.3	24.1 ± 17.2	0.042
INR	2.9 ± 1.8	2.8 ± 1.2	0.371
HBV DNA (log10 IU/mL)	5.6 ± 1.6	5.5 ± 1.6	0.850
HBeAg (positivity rate, %)	152 (54.3)	73 (59.8)	0.303
MELD-based scoring systems			
MELD	28.3 ± 6.2	27.9 ± 6.4	0.508
MELD-Na	39.3 ± 14.2	37.5 ± 13.8	0.218
MELDNa	30.5 ± 5.6	30.0 ± 5.8	0.372
MESO	30.0 ± 22.0	28.3 ± 19.4	0.479
iMELD	60.4 ± 32.0	58.1 ± 31.4	0.510

ANN, artificial neural network; AFP, alpha-fetoprotein; ALT, alanine aminotranferase; AST, aspartate aminotranferase; Hb, hemoglobin; HBeAg, hepatitis B e antigen; iMELD, incorporating serum sodium and age MELD model; INR, international normalized ratio; MELD, the model for end-stage liver disease; MELD-Na, serum sodium and model for MELD; MESO, MELD to serum sodium ratio; PTA, prothrombin activity; TBil, total bilirubin.

CI: 0.823–0.906), compared to that of the MELD-based scoring systems, MELD (AUROC = 0.667, 95% CI: 0.608–0.722, $P < 0.001$), MELD-Na (AUROC = 0.591, 95% CI: 0.531–0.649, $P < 0.001$), MELDNa (AUROC = 0.643, 95% CI: 0.584–0.699, $P < 0.001$), MESO (AUROC = 0.571, 95% CI: 0.510–0.629, $P < 0.001$), iMELD (AUROC = 0.577, 95% CI: 0.517–0.636, $P < 0.001$), respectively (Table S3, Fig. 3a).

When the ANN was finally evaluated in the validation cohort, the AUROC was 0.765, 95% CI: 0.680–0.837, performed better than MELD-based scoring systems, MELD (AUROC = 0.599, 95% CI: 0.507–0.687, $P = 0.006$), MELD-Na (AUROC = 0.563, 95% CI: 0.471–0.653, $P = 0.001$), MELDNa (AUROC = 0.521, 95% CI: 0.429–0.613, $P = 0.006$), MESO (AUROC = 0.571, 95% CI: 0.510–0.629, $P < 0.001$) and iMELD (AUROC = 0.540, 95% CI: 0.447–0.630, $P < 0.001$), respectively (Table S3, Fig. 3b).

Table 3 Univariate analysis of risk factors for 3-month mortality in patients with acute-on-chronic hepatitis B liver failure

	Univariate analysis			
Variables	B	OR	95% CI	P
Sex	0.411	1.508	0.933–2.439	0.094
AFP (μg/L)	−0.001	0.999	0.999–1.000	0.054
ALT (IU/L)	0.000	1.000	0.999–1.000	0.578
AST (IU/L)	0.000	1.000	1.000–1.001	0.251
AST/ALT	0.061	1.063	0.881–1.283	0.523
Albumin (g/L)	−0.018	0.982	0.946–1.019	0.328
Creatinine (mg/dL)	0.330	1.391	0.968–1.997	0.074
Platelet (g/L)	−0.001	0.999	0.995–1.003	0.598
INR	0.144	1.155	0.995–1.340	0.058
HBV DNA (log10 IU/mL)	0.008	1.008	0.887–1.146	0.899
Serum sodium (mM)	−0.107	0.899	0.866–0.932	**<0.001**
TBil (mg/dL)	0.025	1.026	1.005–1.047	**0.015**
Age (years)	0.038	1.039	1.022–1.055	**<0.001**
PTA (%)	−0.046	0.955	0.941–0.970	**<0.001**
Hb (g/L)	−0.016	0.984	0.975–0.993	**<0.001**
HBeAg (positivity rate, %)	0.742	2.101	1.399–3.155	**<0.001**

ACHBLF, acute-on-chronic hepatitis B liver failure; AFP, alpha-fetoprotein; ALT, alanine aminotranferase; AST, aspartate aminotranferase; CI, confidence interval; Hb, haemoglobin; HBeAg, hepatitis B e antigen; INR, international normalized ratio; OR, odds ratio; PTA, prothrombin activity; TBil, total bilirubin.

DISCUSSION

In our study, we firstly established a model to predict 3-month mortality risk of ACHBLF on an individual patient level using an ANN system. The ANN model had been trained in a large population of ACHBLF patients ($n = 280$) and validated in another independent cohort ($n = 122$). AUROCs indicated that the predictive accuracy of the ANN was superior to the MELD-based scoring systems. The better performance of ANN corroborated the opinion that the prediction of 3-month mortality in ACHBLF patients was related to its complex, multidimensional and nonlinear function. The MELD score was initially developed by Malinchoc using a proportional hazard model to obtain the relative risk of the defined variables and then built by a logistic regression [4]. The logistic regression, a generalized linear model used for

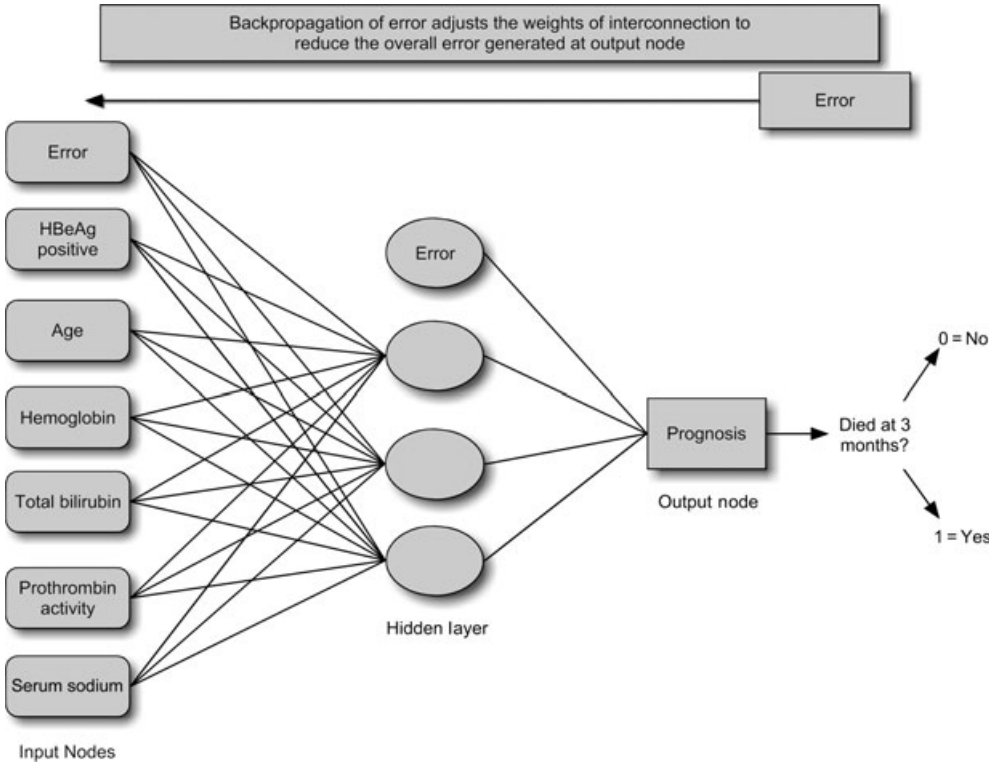

Fig. 2 Schematic representation of the artificial neural network developed to predict 3-month mortality risk in patients with acute-on-chronic hepatitis B liver failure.

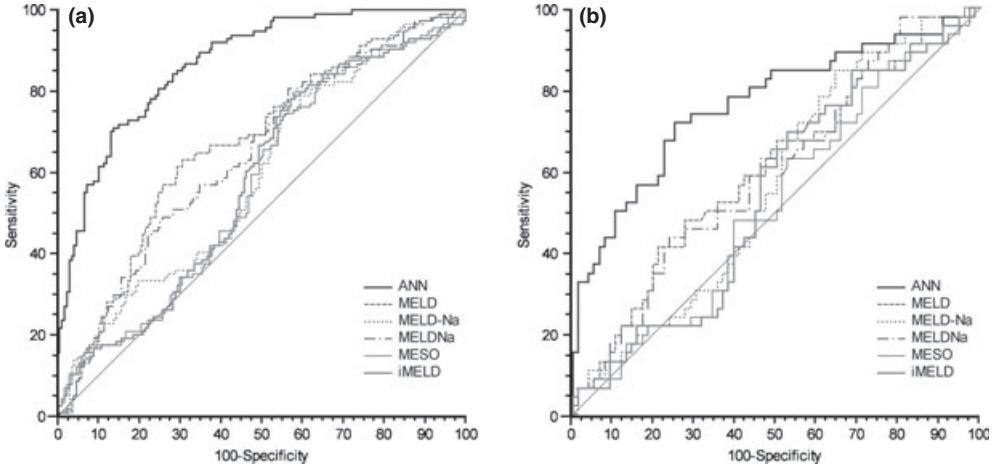

Fig. 3 Receiver operating characteristic curve analysis of the predictive accuracy of the models to predict 3-month mortality of acute-on-chronic hepatitis B liver failure in the (a) training cohort and (b) validation cohort.

binomial regression, was usually used for predicting the probability of an event's occurrence by fitting variables to a logistic curve. Although in recent years, some complemented models were developed based on the original MELD, these were all traditional linear models. Moreover, MELD score or other MELD-based scoring systems were mainly used in the patients of decompensated cirrhosis. ACHBLF represents a complex condition that is significantly different from liver cirrhosis in many aspects, such as short-term and long-term survival rates. ANN is a nonlinear statistical data modelling tool and can be used to model complex relationships.

The present ANN was built and tested on an internal cohort (cross-validation). So there was a possibility that data from other centres might conceive a different prognosis model. However, the ANN could perform an inference process to mitigate errors caused by new data

sets. The ANN can provide a score ranging from 0 (survival) to 1 (death), between which the transitional values were representative of different probabilities of death. The higher the score is, the higher the risk of a 3-month death for ACHBLF patients will be.

In the present study, there were significant differences in PTA, age, serum sodium, TBil, Hb and HBeAg positivity rate between survivors and nonsurvivors. It had been reported that PTA was a risk factor for the mortality of ACHBLF [26]. Furthermore, age, serum sodium and TBil were also risk factors for predicting the prognosis of ACHBLF [5,7,26]. Song *et al.* [27] found that low Hb concentrations were independently associated with regional cerebral oxygen saturation values below 50% in end-stage liver disease patients. Therefore, it was reasonable to believe that Hb might be an important risk factor for ACHBLF. Previous studies also suggested that HBeAg negativity was associated with more severe types of liver disease [28–30].

There are some limitations of this study. First, it is noteworthy that the black-box solution (nonlinear mapping) is a major drawback of neural network models in determining possible internal associations between input and output variables. To solve this problem, further analyses are needed to explain their inference mechanism, such as sensitivity analysis. Second, the study population only came from a single centre and followed up for only 3 months. Multi-centred, prospective studies of larger populations with longer-term follow-up are needed. Finally, dynamic analysis of ACHBLF is more important and meaningful than a single measurement.

In summary, we demonstrated that an ANN was a valid and reliable tool to determine the prognosis of ACHBLF. The ANN we developed was superior to the MELD-based scoring systems in predicting 3-month mortality of ACHBLF.

FINANCIAL SUPPORT

The authors declare that they do not have anything to disclose regarding funding from industries with respect to this manuscript. This work was supported by grants from National Science and Technology Major Project of China (2012ZX10002004), Scientific Research Foundation of Wenzhou, Zhejiang Province, China (H20090014, Y20090269), Health Bureau of Zhejiang Province (2010KYB070), Research Foundation of Education Bureau of Zhejiang Province (Y201009942), Zhejiang Provincial Administration of Traditional Chinese Medicine (2011ZQ015) and Project of New Century 551 Talent Nurturing in Wenzhou, Natural Science Foundation of Shandong Province (ZR2010HQ040), Independent Innovation Foundation of Shandong University (IIFSDU, 2010TS013).

CONFLICTS OF INTEREST

The authors disclose no conflicts.

REFERENCES

1 Sarin SK, Kumar A, Almeida JA, *et al.* Acute-on-chronic liver failure: consensus recommendations of the Asian Pacific Association for the study of the liver (APASL). *Hepatol Int* 2009; 3(1): 269–282.
2 Liu J, Fan D. Hepatitis B in China. *Lancet* 2007; 369(9573): 1582–1583.
3 Arulraj R, Neuberger J. Liver transplantation: filling the gap between supply and demand. *Clin Med* 2011; 11(2): 194–198.
4 Malinchoc M, Kamath PS, Gordon FD, *et al.* A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology* 2000; 31(4): 864–871.
5 Biggins SW, Kim WR, Terrault NA, *et al.* Evidence-based incorporation of serum sodium concentration into MELD. *Gastroenterology* 2006; 130 (6): 1652–1660.
6 Huo TI, Wang YW, Yang YY, *et al.* Model for end-stage liver disease score to serum sodium ratio index as a prognostic predictor and its correlation with portal pressure in patients with liver cirrhosis. *Liver Int* 2007; 27(4): 498–506.
7 Luca A, Angermayr B, Bertolini G, *et al.* An integrated MELD model including serum sodium and age improves the prediction of early mortality in patients with cirrhosis. *Liver Transpl* 2007; 13(8): 1174–1180.
8 Durand F, Valla D. Assessment of the prognosis of cirrhosis: child-Pugh versus MELD. *J Hepatol* 2005; 42Suppl (1): S100–S107.
9 Francoz C, Belghiti J, Castaing D, *et al.* Model for end-stage liver disease exceptions in the context of the French model for end-stage liver disease score-based liver allocation system. *Liver Transpl* 2011; 17(10): 1137–1151.
10 Inaba K, Barmparas G, Resnick S, *et al.* The Model for End-Stage Liver Disease score: an independent prognostic factor of mortality in injured cirrhotic patients. *Arch Surg* 2011; 146(9): 1074–1078.
11 Kamath PS, Wiesner RH, Malinchoc M, *et al.* A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001; 33(2): 464–470.
12 Stewart CA, Malinchoc M, Kim WR, *et al.* Hepatic encephalopathy as a predictor of survival in patients with end-stage liver disease. *Liver Transpl* 2007; 13(10): 1366–1371.
13 Zheng MH, Shi KQ, Fan YC, *et al.* A model to determine 3-month mortality risk in patients with acute-on-chronic hepatitis B liver failure. *Clin Gastroenterol Hepatol* 2011; 9 (4): 351–356.
14 Ho WH, Lee KT, Chen HY, *et al.* Disease-free survival after hepatic resection in hepatocellular carcinoma patients: a prediction approach using artificial neural network. *PLoS ONE* 2012; 7(1): e29179.
15 Cucchetti A, Piscaglia F, Grigioni AD, *et al.* Preoperative prediction of hepatocellular carcinoma tumour

grade and micro-vascular invasion by means of artificial neural network: a pilot study. *J Hepatol* 2010; 52(6): 880–888.
16 Cucchetti A, Vivarelli M, Heaton ND, *et al.* Artificial neural network is superior to MELD in predicting mortality of patients with end-stage liver disease. *Gut* 2007; 56(2): 253–258.
17 Montie JE, Wei JT. Artificial neural networks for prostate carcinoma risk assessment. An overview. *Cancer* 2001; 91(8 Suppl): 1647–1652.
18 Wang D, Wang Q, Shan F, *et al.* 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.
19 Vandenbroucke JP, von ElmE, Altman DG, *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology* 2007; 18(6): 805–835.
20 Wiesner RH, McDiarmid SV, Kamath PS, *et al.* MELD and PELD: application of survival models to liver allocation. *Liver Transpl* 2001; 7(7): 567–580.
21 Kim WR, Biggins SW, Kremers WK, *et al.* Hyponatremia and mortality among patients on the liver-transplant waiting list. *N Engl J Med* 2008; 359(10): 1018–1026.
22 Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995; 346(8982): 1075–1079.
23 Baxt WG. Application of artificial neural networks to clinical medicine. *Lancet* 1995; 346(8983): 1135–1138.
24 Forsstrom JJ, Dalton KJ. Artificial neural networks for decision support in clinical medicine. *Ann Med* 1995; 27(5): 509–517.
25 Dayhoff JE, DeLeo JM. Artificial neural networks: opening the black box. *Cancer* 2001; 91(8 Suppl): 1615–1635.
26 Liu XY, Hu JH, Wang HF. [Analysis of prognostic factors for patients with acute-on-chronic liver failure]. *Zhonghua Gan Zang Bing Za Zhi* 2009; 17(8): 607–610.
27 Song JG, Jeong SM, Shin WJ, *et al.* Laboratory variables associated with low near-infrared cerebral oxygen saturation in icteric patients before liver transplantation surgery. *Anesth Analg* 2011; 112(6): 1347–1352.
28 Chung GE, Kim W, Lee JH, *et al.* Negative hepatitis B envelope antigen predicts intrahepatic recurrence in hepatitis B virus-related hepatocellular carcinoma after ablation therapy. *J Gastroenterol Hepatol* 2011; 26(11): 1638–1645.
29 Wong GL, Wong VW, Choi PC, *et al.* Evaluation of alanine transaminase and hepatitis B virus DNA to predict liver cirrhosis in hepatitis B e antigen-negative chronic hepatitis B using transient elastography. *Am J Gastroenterol* 2008; 103(12): 3071–3081.
30 Ke WM, Li XJ, Yu LN, *et al.* Etiological investigation of fatal liver failure during the course of chronic hepatitis B in southeast China. *J Gastroenterol* 2006; 41(4): 347–351.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Characteristics of the patients with acute-on-chronic hepatitis B liver failure, stratified by mortality.

Table S2. Characteristics of acute-on-chronic hepatitis B liver failure patients in training and validation cohorts, stratified by mortality.

Table S3. ROC analysis of ANN in comparison with MELD-based scoring systems.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.