Am J Cancer Res 2022;12(6):2876-2890
www.ajcr.us /ISSN:2156-6976/ajcr0142402

Original Article

5-year recurrence prediction after hepatocellular carcinoma resection: deep learning vs. Cox regression models

Hon-Yi Shi[1,2,3,4], King-The Lee[1,5], Chong-Chi Chiu[6,7], Jhi-Joung Wang[8,9], Ding-Ping Sun[10], Hao-Hsien Lee[11]

[1]Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, Kaohsiung 80708, Taiwan; [2]Department of Business Management, National Sun Yat-sen University, Kaohsiung 80420, Taiwan; [3]Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung 80708, Taiwan; [4]Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan; [5]Hepatobiliary-Pancreatic Surgery, Park One International Hospital, Kaohsiung 81357, Taiwan; [6]Department of General Surgery, E-Da Cancer Hospital, Kaohsiung 82445, Taiwan; [7]School of Medicine, College of Medicine, I-Shou University, Kaohsiung 82445, Taiwan; [8]Department of Medical Research, Chi Mei Medical Center, Yongkang, Tainan 71004, Taiwan; [9]Allied AI Biomed Center, Southern Taiwan University of Science and Technology, Tainan 71005, Taiwan; [10]Department of Surgery, Chi Mei Medical Center, Yongkang, Tainan 71004, Taiwan; [11]Department of Surgery, Chi Mei Medical Center, Liouying, Tainan 73658, Taiwan

Received February 16, 2022; Accepted May 30, 2022; Epub June 15, 2022; Published June 30, 2022

Abstract: Deep learning algorithms have yet to be used for predicting clinical prognosis after cancer surgery. Therefore, this study compared performance indices and permutation importance of potential confounders in three models for predicting 5-year recurrence after hepatocellular carcinoma (HCC) resection: a deep-learning deep neural network (DNN) model, a recurrent neural network (RNN) model, and a Cox proportional hazard (CPH) regression model. Data for 725 patients who had received HCC resection at three medical centers in southern Taiwan between April, 2011, and December, 2015, were randomly divided into three datasets: a training dataset containing data for 507 subjects was used for model development, a testing dataset containing data for 109 subjects was used for internal validation, and a validating dataset containing data for 109 subjects was used for external validation. Feature importance analysis was also performed to identify potential predictors of recurrence after HCC resection. Univariate Cox proportional hazards regression analyses were performed to identify potential significant predictors of 5-year recurrence after HCC resection, which were included in the forecasting models ($P < 0.05$). All performance indices for the DNN model were significantly higher than those for the RNN model and the conventional CPH model ($P < 0.001$). The most important potential predictor of 5-year recurrence after HCC resection was surgeon volume followed by, in order of importance, hospital volume, preoperative Beck Depression Scale score, preoperative Beck Anxiety Scale score, co-residence with family, tumor stage, and tumor size. The feature importance analysis performed to investigate interpretability in this study elucidated the potential use of deep learning models for predicting recurrence after HCC resection and for identifying predictors of recurrence. Further experiments using the proposed DNN model would clarify its potential uses for developing, promoting, and improving health policies for treating HCC patients after surgery.

Keywords: Hepatocellular carcinoma, resection, deep learning, recurrence, feature importance

Introduction

Hepatocellular carcinoma (HCC) remains the fourth leading cause of cancer-related death in the world. Liver resection is the mainstay of curative treatment for HCC, but HCC recurrence after resection remains high; 60% to 70% of patients who undergo HCC resection develop recurrence within 5 years after surgery [1-3]. In patients who have undergone HCC resection, recurrence is significantly associated with demographic attributes, aggressive tumor pathology, surgeon volume, hospital volume, and preoperative quality of life [1-3]. Therefore, accurately predicting recurrence after HCC resection is critical for effective delivery of precision medicine. The capability to predict recurrence in individual patients is also essential because it enables clinicians to gauge treatment outcomes.

Recently developed deep-learning algorithms based on multilayer perceptrons use clinical pathological data to predict cancer surgery outcomes [4, 5]. The utility of deep learning algorithms has been examined in various translational and clinical medicine settings [6-11]. However, literature on their use for predicting cancer recurrence is relatively sparse. For disease recurrence analysis, deep-learning algorithms, a class of machine learning models, can use nonlinear risk functions for automatically modeling recurrence risk and for predicting individual medical outcomes based on learned representations. Recent advances in artificial intelligence, including deep learning algorithms, have also been used for outcome prediction in the medical field [4-11]. We investigated whether the baseline models in our study, which were either recurrent neural network (RNN) or deep neural network (DNN) models, improved efficiency via neural network condensation without sacrificing prediction accuracy. Therefore, we were interested in identifying the input features that improved the accuracy of RNN models and DNN models in using electronic medical data for continuous assessment of 5-year recurrence risk after HCC resection. However, deep learning algorithms have yet to be used for predicting clinical prognosis after cancer surgery. The main reason is the high complexity of prediction algorithms relative to diagnostic algorithms, e.g., prediction algorithms require more numerous parameters. The selective studies regarding to recurrence after HCC surgery are summarized in **Table 1** [6-11]. These studies agree that, for predicting disease-free recurrence of HCC and overall recurrence, multi-dimensional artificial intelligence analysis outperforms Cox proportional hazards analysis. However, no studies have compared deep learning algorithms and Cox proportional hazards regression models in terms of accuracy in predicting post-operative recurrence after HCC resection.

Therefore, the aim of the long-term prospective study were to evaluate the performance indices and to conduct the permutation importance of potential confounders by using the deep-learning deep neural networks (DNNs) and recurrent neural networks (RNNs), and a Cox proportional hazard (CPH) regression model for predicting 5-year recurrence in a large population of HCC patients who have undergone resection.

Materials and methods

Subjects and study design

The subjects of this study were HCC patients who had received surgical resection performed at one of three southern Taiwan medical centers between April, 2011, and December, 2015. Inclusion criteria were the following: (1) a histologic or a combined radiographic and laboratory diagnosis of HCC; (2) ability to communicate in Chinese or Taiwanese; and (3) agreement to participate in a questionnaire survey performed in the hospital ward. Major exclusion criteria included loss to follow up or death during the study period, which might have interfered with the analysis in this study. **Figure 1** shows that, during the sample selection period, 819 subjects were eligible for participation. Of these, 94 were excluded due to loss to follow up. Therefore, 725 subjects participated in the preoperative (baseline) and postoperative 5-year assessments in this study. The institutional review boards at all participating hospitals approved the study protocol (KMUH-IRB-20110002). Additionally, all participants gave informed consent before they enrolled in this study.

Instruments and potential confounders

Research assistants collected demographic, clinical, and institutional data by reviewing questionnaire survey results and medical records. Data collection included gender, age, marital status, education, co-residence with the family, hepatitis B, hepatitis C, body mass index (BMI), Charlson comorbidity index (CCI), tumor size, tumor stage, American Society of Anesthesiologists score, average length of stay, 30-day readmission after hospital discharge, surgeon volume, hospital volume, and preoperative Beck Anxiety Inventory (BAI) score and Beck Depression Inventory (BDI) score. The BAI and BDI were used to evaluate anxiety and depression in the HCC patients before surgery [12, 13]. Each scale includes 21 self-evaluated items, each of which is scored from 1 to 4 point(s). A high BAI score or BDI indicates high severity of anxiety or depression, respec-

Table 1A. Selective studies in predicting recurrence after hepatocellular carcinoma (HCC) surgery

Authors (country)	No. of subjects	Deep learning algorithms	Major findings
Tang, et al. (USA) [6]	197 patients with confirmed locally advanced oesophageal squamous cell carcinoma (SCC)	1. Radiomics model 2. Clinical model 3. Integration of radiomics and clinical models	1. The area under the curve (AUC) of integration of radiomics and clinical models was better than that of radiomics or clinical model for the training cohort (0.821 vs. 0.754 or 0.679, respectively) and for the validation cohort (0.809 vs. 0.646 or 0.658, respectively). 2. Integrated model of radiomics and clinical features showed good performance in predicting early recurrence of locally advanced oesophageal SCC for both the training and validation cohorts (accuracy = 0.730 and 0.733, and F-1 score = 0.730 and 0.778, respectively).
Nam et al. (South Korea) [7]	563 patients who underwent liver transplantation for HCC	1. MoRAL-artificial intelligence (AI) model 2. Milan criteria model 3. MoRAL model 4. University of California San Francisco (UCSF) model 5. Up-to-seven model 6. Kyoto criteria model	The MoRAL-AI showed significantly better discrimination function (c-index = 0.75) than the Milan (c-index = 0.64), MoRAL (c-index = 0.69), UCSF (c-index = 0.62), up-to-seven (c-index = 0.50), and Kyoto (c-index = 0.50) criteria (all $P < 0.001$).
Ji et al. (China) [8]	470 patients with solitary HCC	1. Maximum relevance minimum redundancy (MRMR) 2. Random survival forest (RSF) 3. Least absolute shrinkage and selection operator (LASSO) Cox regression	1. A three-feature signature that demonstrated favorable prediction of HCC recurrence across all datasets, with C-index of 0.633-0.699. 2. The two models showed superior prognostic performance, with C-index of 0.733-0.801 and integrated Brier score of 0.147-0.165, compared with rival models without radiomics and widely used staging systems (all $P < 0.05$).

Table 1B. Selective studies in predicting recurrence after cancer surgery

Authors (country)	Number of cases	Deep learning algorithms	Major findings
Liu, et al. (Taiwan) [9]	97 patients with HCC	1. Multi-dimensional artificial intelligence analysis 2. Alpha fetoprotein level 3. Cox proportional hazards analysis	1. The multi-dimensional artificial intelligence analysis can better predict early recurrence (area under the curve = 0.917, sensitivity = 81.8%, specificity = 90.5%), compared to alpha fetoprotein level (area under the curve = 0.595, sensitivity = 68.2%, specificity = 47.6%). 2. The multi-dimensional artificial intelligence analysis also showed a better performance in prediction of disease-free survival and overall survival than a Cox proportional hazards analysis.
Mai et al. (China) [10]	903 patients with HCC	1. Artificial neural network (ANN) model 2. Cox proportional hazards (CPH) model	The subsequent receiver-operating characteristic (ROC) analysis revealed that the ANN model predicted post-hepatectomy early recurrence (PHER) better than the CPH model (ANN: 0.753 versus CPH: 0.733, $P < 0.05$) as well as the eight prognostic factors individually (ANN: 0.753 versus corresponding ROCs: 0.534-0.624, $P < 0.05$) for all comparisons.
Song et al. (China) [11]	601 patients with HCC	1. Deep learning model 2. Radiomics model 3. Deep learning combined with clinical parameters model	1. To predict the presence of microvascular invasion (MVI), the deep learning model based only on images achieved an area under curve (AUC) of 0.915 in the testing cohort as compared to the radiomics model with an AUC of 0.731. 2. The deep learning combined with clinical parameters model yielded the best predictive performance with an AUC of 0.931.

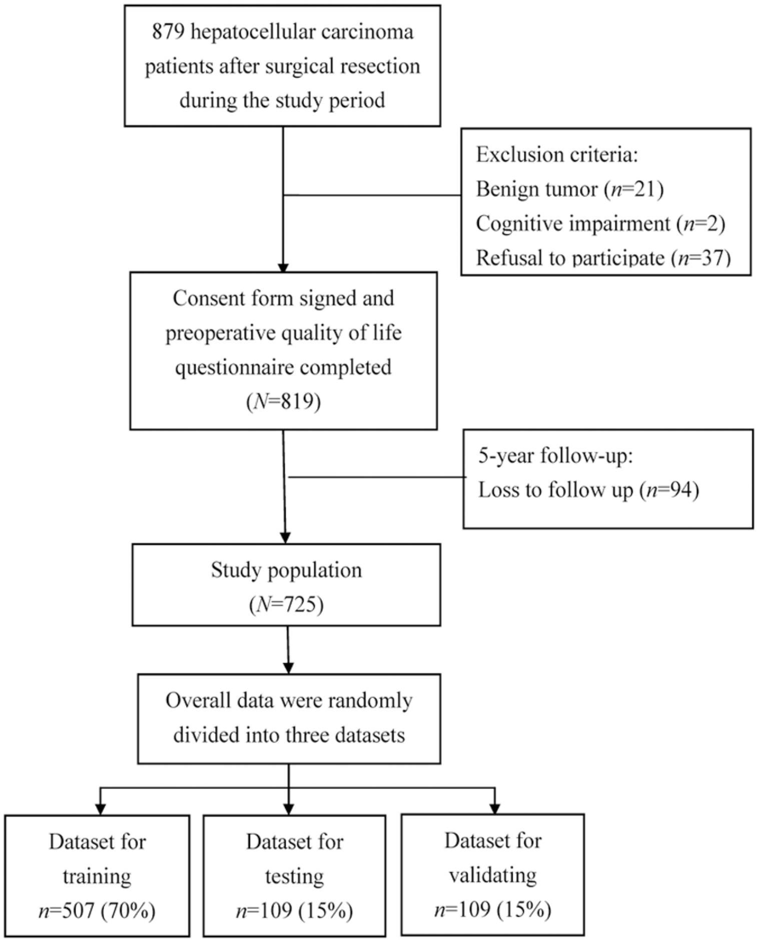

Figure 1. Flowchart of the study procedure.

tively. Thus, a high score before and after therapy suggests that the therapy was ineffective. The Chinese versions of the BDI and BAI used in this study have been validated and widely used in both clinical practice and research [14]. The surgeon volume and hospital volume data included in the dataset were sorted by total annual surgical procedures, and each procedure was assigned a unique identification code. In performance comparisons of the DNN, RNN, and CPH models, the above variables were the independent variables, and 5-year recurrence after HCC resection was the dependent variable.

Development of the deep learning algorithms

The neural network type most commonly used for real-world applications is the feedforward neural network, in which neurons are organized into distinct layers in order of topology. The first layer is the input layer, the last layer is the output layer, and the other layers are the hidden layers. This study defined a DNN as a feedforward network with more than one hidden layer. The DNN model was trained using back-propagation [15]. First, the transfer learning framework was used to train only the top layers that were randomly initialized by freezing all convolutional layers. The model was then trained using several epochs of data from our dataset, and validation loss was monitored by applying early stop criteria. The typical feature of the RNN architecture is a cyclic connection, which enables the RNN to update its current state based on past states and current input data [16]. Finally, grid-search method was used to optimize the hyperparameters and architectures of the deep learning DNN and RNN algorithms, and the numbers of epochs were determined by performing the tuning procedure described in **Table 2**.

Statistical analysis

The unit of analysis in this study was the individual patient with HCC after resection. The study characteristics were reported as mean ± standard deviation (SD) and median interquartile range (IQR) for continuous variables and as frequency (%) for categorical variables. Cox proportional hazards regression was performed to investigate hazard ratios (HRs) and 95% confidence intervals (95% CI) and to construct a model for predicting 5-year recurrence after HCC resection. Recurrence time was defined as the interval between the first HCC diagnosis and the first HCC recurrence.

Statistical analysis was performed in the following steps. In the first step, univariate Cox proportional hazards regression analyses were performed to identify significant ($P < 0.05$) predictors of 5-year recurrence after HCC resec-

Table 2. Parameters for classification models

Parameters	Deep neural networks	Recurrent neural networks
No. of hidden layers	4	4
No. of neuron in each hidden layers	(64, 64, 128, 256)	(64, 64, 128, 256)
Activation functions in each layer	Rectified linear unit (ReLU) in hidden layers and sigmoid on output layer	
Loss function	Binary cross entropy	
Optimizer	Adaptive moments estimation (Adam) with 0.001 learning rate	
No. of Epochs	100	100
Dropout layers for regularization	20% dropout layer after second hidden layer and 10% after third hidden layers	

tion, which were then included in the forecasting models. In the second step, data for the study cohort of 725 subjects were randomly divided into three datasets: a training dataset containing data for 507 subjects was used for model development, a testing dataset containing data for 109 subjects was used for internal validation, and a validating dataset containing data for 109 subjects was used for external validation. The independent variables fitted to the forecasting models were significant predictors of 5-year recurrence after HCC resection, and the dependent variable was 5-year recurrence after HCC resection. After model training, model outputs were collected for each testing dataset. In the third step, 1,000 pairs of forecasting models with 95% confidence intervals were compared in terms of indices of performance in predicting 5-year recurrence after HCC resection. Independent t test was used to identify performance indices that significantly differed among the three models. Indices used for performance comparisons included accuracy, precision, recall, F-1 score, and area under the receiver operating characteristic (AUROC) curve. In the fourth and final step, permutation importance was calculated by using an algorithm to obtain an importance score for each confounder variable in the dataset [17]. Permutation importance was defined as the model score when the value of a single confounder was randomly changed. This technique has two advantages: first, it is model-agnostic; second, it can be performed repeatedly with different feature permutations.

The scikit-learn 0.21.2 implementation for Python (version 3.7.6; Python Software Foundation, Wilmington, Delaware, USA) was used to run the deep learning DNN and RNN models. The CPH model was computed with the Lifelines v0.22.2 implementation for Python v3.7.6 and double-checked with JMP10.0 (SAS Institute Inc., Cary, NC). All statistical tests were two-sided; a p value less than 0.05 was considered statistically significant.

Results

Study characteristics

Table 3 shows that, during the study period, 5-year recurrence after HCC resection was 50.8%. In patients with HCC recurrence after undergoing surgery, mean age was 61.4 years (SD, 10.7 years), median age was 62 years (IQR 55-69 years), and most (70.4%) patients were male. The univariate analysis revealed that gender, age, marital status, education, co-residence with family, hepatitis B, hepatitis C, BMI, CCI, tumor size, tumor stage, American Society of Anesthesiologists score, average length of stay, 30-day readmission after hospital discharge, surgeon volume, hospital volume, and preoperative BAI and BDI scores are significantly associated with 5-year recurrence after HCC resection ($P < 0.05$) (**Table 4**).

Comparison of forecasting models

The study characteristics did not significantly differ between the training and testing datasets (**Table 5**); therefore, samples were compared between the training and testing datasets to increase reliability of the validation results. In **Figure 2**, the performance indices for the training dataset indicate that the DNN model had significantly higher accuracy, precision, recall, F1 score, and AUROC values compared to the RNN and CPH models ($P < 0.001$).

Permutation importance of potential confounders in the DNN model

Figure 3 shows the results of permutation importance of confounders in the DNN model. Surgeon volume and hospital volume were the

Table 3. Patient characteristics (N = 725)

Variables	Mean ± standard deviation/median [IQR]	N (%)
Gender		
Male		510 (70.4%)
Female		215 (29.6%)
Age, years	61.4±10.7/62 [55-69]	
Marital status		
Divorced or widowed		96 (13.3%)
Married		629 (86.7%)
Education, years	12.5±4.9/12 [9-16]	
Co-residence with the family		
Yes		636 (87.7%)
No		89 (12.3%)
Hepatitis B		
Yes		227 (31.3%)
No		498 (68.7%)
Hepatitis C		
Yes		138 (19.0%)
No		587 (81.0%)
Body mass index, kg/m^2	24.7±3.5/24.5 [22.4-26.6]	
Charlson Comorbidity Index, score	1.2±1/1 [0-2]	
Tumor size, cm	4.7±3.4/3.8 [2.5-5.9]	
Tumor stage		
Stage I		349 (48.1)
Stage II		216 (29.8)
Stage III		102 (14.0)
Stage IV		58 (8.1)
American Society of Anesthesiologists, score		
I		181 (25.0%)
II		275 (37.9%)
≥III		269 (37.1%)
Average length of stay, days	10.9±7.9/9 [7-11]	
30-day readmission after discharge		
Yes		91 (12.5%)
No		634 (87.5%)
Surgeon volume, cases/year	5.3±2.4/7 [1-14]	
Hospital volume, cases/year	21.4±18.5/57 [23-97]	
Preoperative BDI score	4.8±4/5 [2-5]	
Preoperative BAI score	2±2.8/2 [0-3]	
5-year recurrence after surgery		
Yes		368 (50.8%)
No		357 (49.2%)

IQR, interquartile range; BDI, Beck Depression Inventory; BAI, Beck Anxiety Inventory.

most important confounders when predicting 5-year recurrence after HCC recurrence followed by preoperative BDI score, preoperative BAI score, co-residence with family, tumor stage, and tumor size.

Sensitivity analysis

Next, performance indices obtained for the testing dataset in internal validation and for the validating dataset in external validation

Table 4. Univariate Cox regression analysis of 5-year recurrence after HCC resection (N = 725)

Variables	HR (95% CI)	*P* value
Gender		
Male vs. female	1.19 (0.65, 1.73)	< 0.001
Age, years	0.30 (0.17, 0.51)	< 0.001
Marital status		
Married vs. divorced or widowed	0.29 (0.22, 0.39)	< 0.001
Education, years	0.91 (0.88, 0.93)	< 0.001
Co-residence with the family		
Yes vs. no	0.32 (0.25, 0.43)	< 0.001
Hepatitis B		
Yes vs. no	2.56 (1.44, 4.57)	< 0.001
Hepatitis C		
Yes vs. no	2.87 (1.63, 5.08)	< 0.001
Body mass index, kg/m^2	0.96 (0.94, 0.97)	< 0.001
Charlson comorbidity index, score	1.23 (1.13, 1.35)	< 0.001
Tumor size, cm	1.17 (1.11, 1.23)	< 0.001
Tumor stage		
Stage II vs. stage I	4.93 (3.22, 7.52)	< 0.001
Stage III vs. stage I	12.57 (4.82, 32.78)	< 0.001
Stage IV vs. stage I	54.23 (13.32, 220.91)	< 0.001
American Society of Anesthesiologists, score		
II vs. I	4.05 (2.51, 6.54)	< 0.001
≥III vs. I	1.79 (1.24, 2.60)	< 0.002
Average length of stay, days	1.10 (1.07, 1.13)	< 0.001
30-day readmission after discharge		
Yes vs. no	2.59 (1.82, 3.39)	< 0.001
Surgeon volume, cases/year	0.13 (0.09, 0.17)	< 0.001
Hospital volume, cases/year	0.11 (0.06, 0.17)	< 0.001
Preoperative BDI score	1.13 (1.09, 1.17)	< 0.001
Preoperative BAI score	1.16 (1.10, 1.22)	< 0.001

HCC, hepatocellular carcinoma; HR, hazard ratio; BDI, Beck Depression Inventory; BAI, Beck Anxiety Inventory; CI, confidence interval.

were compared among the forecasting models (**Table 6**). **Table 6** shows that, for both the testing and validating datasets, the DNN model had better performance indices compared to the RNN and CPH models.

Discussion

To the best of our knowledge, this study is the first to use deep-learning DNN and RNN models to investigate 5-year recurrence after HCC resection. Indices of performance in predicting 5-year recurrence after HCC resection were compared among the DNN, RNN, and CPH models. For a given set of demographic, clinical, and institutional inputs, the DNN model clearly had superior performance indices compared to the other two models. Notably, our long-term prospective study collected longitudinal registry data from three different medical institutions, which provided a real-world depiction of current surgical treatment for HCC patients. In contrast, previous works have used data from a single medical center [6-11]. Moreover, using registry data obtained from three medical centers mitigated the potential for referral bias or bias caused by analyzing the practices of a single surgeon or a single institution.

In this long-term prospective study, our analysis demonstrated that deep-learning DNN and RNN models are superior to conventional CPH model in predicting HCC recurrence after re-

Table 5. Comparisons of study characteristics between training dataset and testing dataset after HCC resection

	Training dataset (N = 507)	Testing dataset (N = 109)	*P* value
Gender			0.949
Male	359 (70.9%)	77 (70.5%)	
Female	148 (29.1%)	32 (29.5%)	
Age, years†	61.1±10.5	63.3±12.0	0.134
Marital status			0.347
Divorced or widowed	73 (14.3%)	11 (10.3%)	
Married	434 (85.7%)	98 (89.7%)	
Education, years†	12.5±5.0	11.7±5.3	0.200
Co-residence with family			0.291
Yes	447 (88.2%)	91 (83.3%)	
No	60 (11.8%)	18 (16.7%)	
Hepatitis B			0.389
Yes	168 (33.2%)	31 (28.2%)	
No	489 (66.8%)	78 (71.8%)	
Hepatitis C			0.718
Yes	93 (18.4%)	18 (16.7%)	
No	414 (81.6%)	91 (83.3%)	
Body mass index, kg/m^2†	24.8±3.6	24.5±3.3	0.359
Charlson comorbidity index, score	1.1±1.0	1.2±1.0	0.547
Tumor size, cm†	4.4±3.3	4.1±3.0	0.472
Tumor stage			0.233
Stage I	235 (46.4%)	61 (56.4%)	
Stage II	162 (31.9%)	25 (23.1%)	
Stage III	69 (13.7%)	17 (15.4%)	
Stage IV	41 (8.0%)	6 (5.1%)	
American Society of Anesthesiologists, score			0.752
I	131 (25.8%)	28 (25.6%)	
II	184 (36.3%)	43 (39.7%)	
≥III	192 (37.9%)	38 (34.6%)	
Average length of stay, days†	11.1±7.8	10.9±9.3	0.794
30-day readmission after discharge			0.290
Yes	65 (12.9%)	10 (9.0%)	
No	442 (87.1%)	99 (91.0%)	
Surgeon volume, cases/year†	5.2±4.7	5.4±4.5	0.498
Hospital volume, cases/year†	22.1±12.8	21.0±11.9	0.378
Preoperative BDI score†	4.7±4.0	5.1±4.4	0.542
Preoperative BAI score†	2.1±2.9	2±2.7	0.788

†Values are expressed as means ± standard deviation or N (%). HCC, hepatocellular carcinoma; BDI, Beck Depression Inventory; BAI, Beck Anxiety Inventory.

section. A previous literature indicates that an increasing number of studies are integrating deep-learning models in analytic approaches used in oncologic research [4]. Most of these studies are related to either diagnostic work-up, such as radiographic image analysis and cytopathologic interpretation, or genomic/molecular analysis for biomarker discovery; to date, studies that have used deep-learning models for predicting recurrence in oncology patients remain limited.

For analyzing recurrence after HCC resection, DNN models have three main strengths. First,

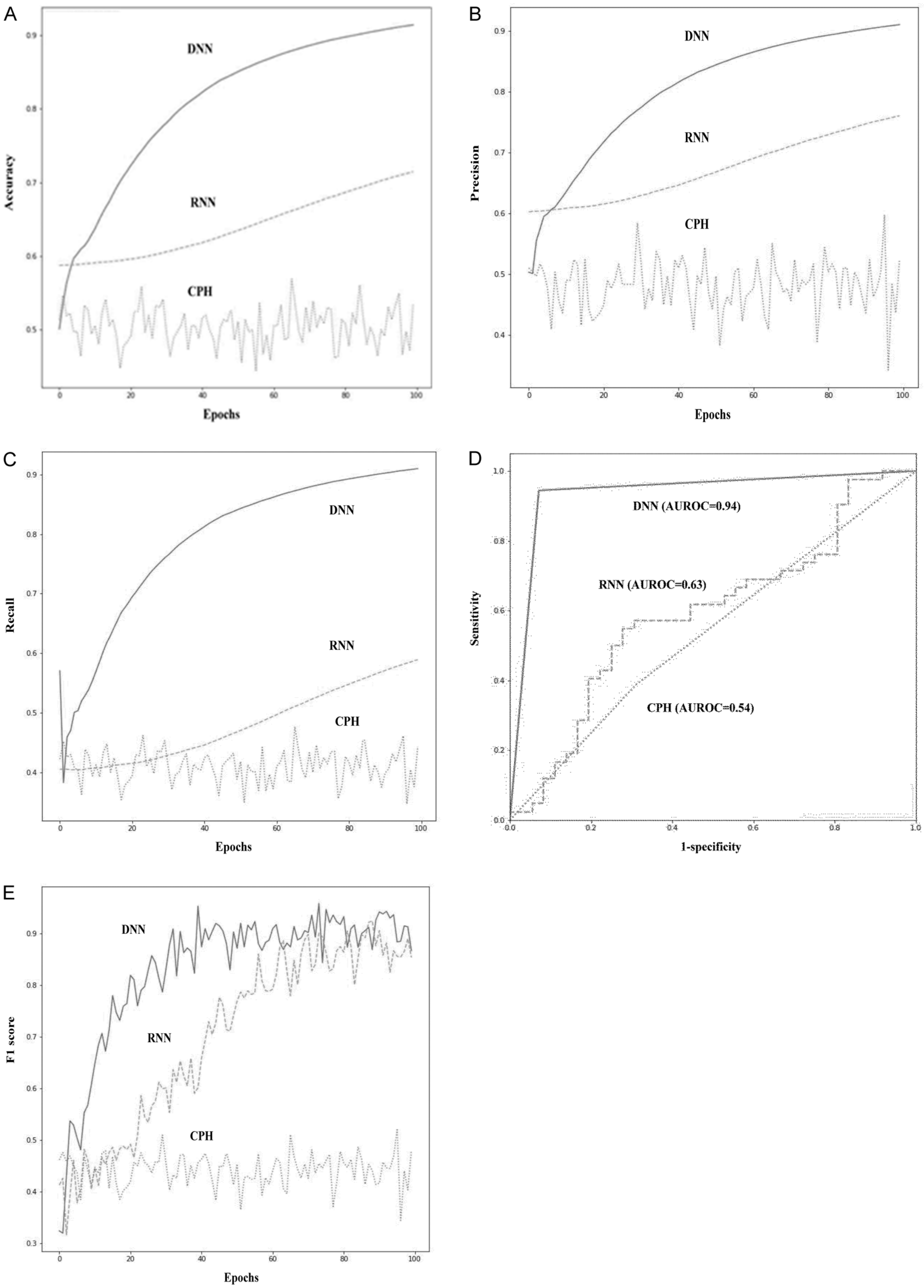

Figure 2. Comparison of (A) accuracy, (B) precision, (C) recall, (D) F1 score, and (E) area under the receiver operating characteristics (AUROC) curve among the deep neural network (DNN), recurrent neural network (RNN), and Cox proportional hazard regression (CPH) models for predicting 5-year recurrence after hepatocellular carcinoma (HCC) resection in the training dataset.

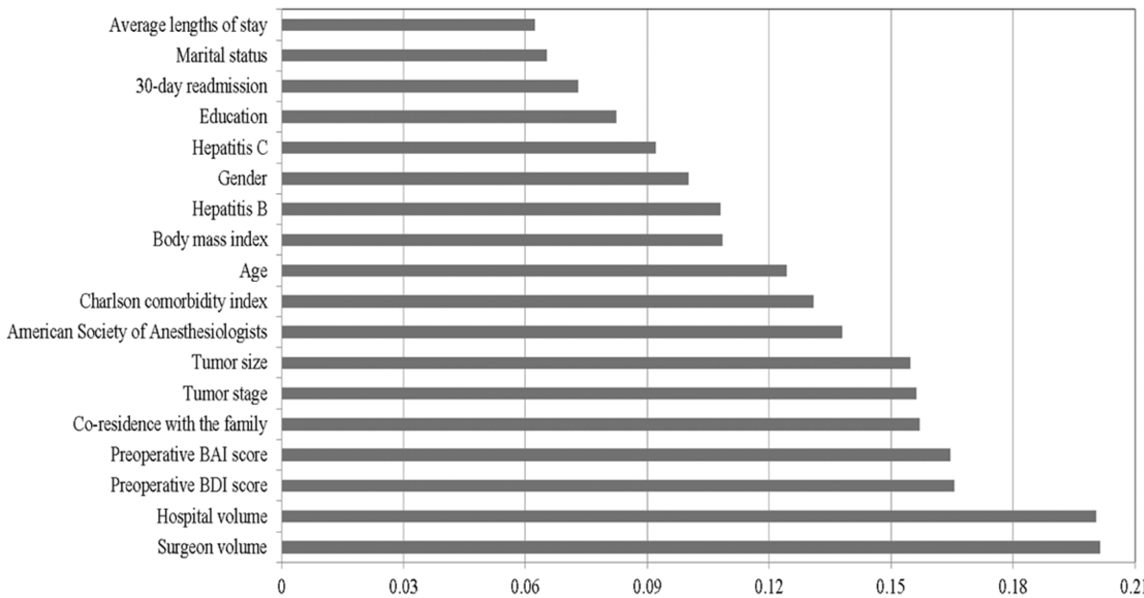

Figure 3. Permutation importance of the potential confounders in the deep neural network (DNN) model.

Table 6. Comparison of performance indices between deep learning and Cox regression models used to predict 5-year recurrence after HCC resection

Models	Accuracy	Precision	Recall	F1 Score	AUROC
Testing dataset (N = 109)					
Deep neural network	84.39%	83.46%	82.11%	79.92%	0.93
Recurrent neural network	60.63%	62.51%	51.14%	63.57%	0.83
Cox regression	53.57%	50.84%	34.78%	40.64%	0.52
Validating dataset (N = 109)					
Deep neural network	84.74%	84.60%	83.56%	79.17%	0.93
Recurrent neural network	59.20%	71.73%	71.78%	61.96%	0.83
Cox regression	50.73%	68.36%	39.23%	49.50%	0.54

HCC, hepatocellular carcinoma; AUROC, area under the receiver operating characteristic.

as discussed earlier, use of DNN models enables a superior fit to variables with nonlinear relationships, which is applicable when examining real-life variables. Notably, previous works have demonstrated that numerous clinical-laboratory variables are non-linearly associated with medical outcomes and have implied that deep-learning models may be more appropriate than linear regression models in terms of predicting risk of various medical outcomes [6-11]. Second, deep-learning models can automatically learn feature representations from raw clinical data without explicit feature engineering; additionally, nonlinear risk functions can be used for fitting deep-learning models to censored data for recurrence after resection [4, 7]. That is, deep-learning models excel at identifying nonlinear relationships among data, and they easily accommodate censoring of recurrence data after resection. Third, this study suggests that DNN and RNN models outperform the CPH model when large feature sets are used. Since a deep-learning model can learn feature representation, perhaps the most important strength of the deep-learning model is its capability to analyze large feature sets, which is particularly beneficial in biomedical research since including numerous variables in conventional linear regression models can cause overfitting.

In line with previous studies, treatment at a high-volume hospital or by a high volume surgeon was significantly associated with good medical outcomes [18-20]. A previous study of a nationwide population provided additional insight into the combined effects of surgeon volume and hospital volume on survival after hepatic resection. The authors concluded that treatment by a high-volume surgeon and treatment in a high-volume hospital were the two important predictors of short-term survival after hepatic resection [18]. Notably, they defined high-volume hospitals as those that performed more than 245 hepatic resection procedures annually and defined high-volume surgeons as those who performed more than 59 hepatic resection procedures annually. Another study reviewed proposed guidelines for hepatobiliary surgery and addressed the effects of volume on outcomes of hepatobiliary surgery. The authors recommended the establishment of an expert panel to draft preparatory and preliminary agreements on the clarification and standardization of technical terminology used to discuss HCC resection and indicators for HCC resection [19]. Lu et al. retrospectively analyzed 23,107 major hepatectomies for HCC patients and concluded that surgeon volume and hospital volume are significant independent predictors of postoperative recurrence ($P < 0.001$) [20]. The best explanation for this finding is that 'practice makes perfect', i.e., surgeons who perform a high volume of HCC resection procedures are the most likely to achieve a high level of surgical skill, which then reduces the postoperative rate of HCC recurrence in their patients. Likewise, hospitals that perform a high volume of HCC resection procedures have the lowest recurrence rates because they have the most opportunities and incentive to establish and implement an effective system for preventing recurrence throughout a given episode of inpatient care. Another possible explanation is that high volume hospitals have the most incentive and the most opportunities to develop standard procedures for care and after complex surgical procedures such as HCC resection.

Regarding use of both preoperative BDI score and preoperative BAI score as predictors of cancer recurrence, Wang et al. performed a meta-analysis of patient data from a selection of 17 eligible cohort studies involving 282,203 breast cancer patients to investigate whether baseline preoperative BAI and BDI scores are prognostic predictors of cancer recurrence [21]. Their meta-analysis confirmed that depression, anxiety, and the combination of depression and anxiety are associated with recurrence in patients with breast cancer. Several factors may explain the increased recurrence of cancer in patients with depression and anxiety, including unhealthy lifestyle factors, treatment nonadherence, and biological mechanisms. A recent multicenter, cross-sectional study of 996 patients diagnosed with mixed tumors also showed that patients newly diagnosed with cancer tend to have high anxiety about cancer recurrence [22]. The study also found that fear of cancer recurrence was significantly associated with financial burden, receipt of chemotherapy, severe illness in childhood, and emotional disturbances (anxiety and depression). In Chen et al., a nationwide population-based cohort study was performed to investigate the association between depressive disorders and tumor recurrence risk in patients who had received curative surgery for breast cancer [23]. The cohort with depressive disorder had a significantly higher risk of postoperative recurrence compared to the cohort without depressive disorder (HR = 1.373, 95% CI 1.09-1.72, $P = 0.005$). A possible explanation is that impaired endocrine function and impaired hypothalamic-pituitary-adrenal function in patients with depressive disorder may contribute to increased risk of recurrence after breast cancer surgery.

Su et al. administered the Mini International Neuropsychiatric Interview in breast cancer patients to investigate whether family support is associated with major depressive disorder [24]. The study found that family support, which has a well-established positive association with breast cancer outcome, also has a negative association with depression risk. Another prospective 23-year study of 373 subjects similarly reported that, compared to patients with weak family support, those who had strong family support tended to have lower severity of depression at baseline and more rapid improvement in depression [25]. Family caregivers may be ill prepared to assume responsibilities such as providing family members with information on the disease and treatment or providing instruction in technical and care skills [26]. Moreover, caregiving by family members must be balanced against already established roles and role responsibilities. Additionally, family members who assume a caregiving role may have their own emotional responses to the diagnosis and prognosis of the patient and may require coaching and emotional support. As previously noted, however, further studies are needed to explore the influence of family support on HCC recurrence after resection.

In agreement with previous works [27, 28], the present study found that advanced cancer stage was significantly associated with recurrence within 5 years after HCC resection. The patients analyzed in this study included 250 patients (48.1%) in tumor stage I, 155 (29.8%) in tumor stage II, 73 (14.0%) in tumor stage III, and 42 (8.1%) in tumor stage IV. Early diagnosis of HCC and curative retreatment are likely to help prevent recurrence. After surgery, HCC patients are often burdened by multiple cancer-related comorbidities that increase their risk of poor postoperative outcomes, including complications, a long hospital stay, a short survival time, and high treatment costs. Zhang et al. reported that, after curative resection, independent predictors of HCC recurrence included tumor size (> 5 cm vs. < 5 cm) (odds ratio, OR = 1.45, 95% CI 1.05-2.02) and tumor-node-metastasis (TNM) stage (III-IV vs. I-II) (OR = 1.36, 95% CI 1.02-1.81) [27]. Kumar et al. also pointed out that T-stage and increased tumor size were significantly associated with early postoperative recurrence ($P < 0.001$) [28].

For further validation of the significant association observed between potential predictors and 5-year recurrence after HCC surgery, **Table 7** lists selected studies that have identified significant predictors for 5-year recurrence after HCC surgery [29-33]. As in these previous works, our study demonstrated that low surgeon volume, low hospital volume, high preoperative BDI score, high preoperative BAI score, no co-residence with the family, high tumor stage, and large tumor size are significantly associated with 5-year recurrence after HCC surgery ($P < 0.001$).

This large, long-term, and prospective observational cohort study of HCC patients in Taiwan analyzed data from patients treated at three medical centers. The accuracy of the deep-learning DNN model developed in this study was superior to that of the RNN and CPH models in predicting 5-year recurrence after HCC resection. Three implications of this study are noted. First, the proposed DNN model may be useful for guiding the clinical care of HCC patients after resection. Second, healthcare administrators and managers at medical institutions should facilitate the prompt establishment of surgical teams designed to meet the specific needs of individual patients. Third, the Taiwan National Health Insurance Administration should include both surgical volumes and emotional outcomes in its guidelines for clinical treatment of HCC in order to achieve a broad nationwide improvement in care for these patients. However, further studies are needed to confirm the clinical relevance of the proposed DNN model in terms of its efficacy in predicting prognosis and optimizing medical management for HCC patients after resection.

This study has several limitations inherent in a large database analysis. First, although we likely examined one of the largest sample sizes among studies of this nature, the relatively small total number of patients in this study limited the reliability of the results obtained by the deep-learning models. Second, the analysis was limited to 5-year recurrence after HCC resection, which reduced the subset of HCC patients with surgery in which the DNN model is clinically applicable. Third, this study only compared DNN, RNN and CPH models. Further studies are needed to compare other forecasting models. Finally, this study did not compare other outcomes, e.g., patient-reported quality of life, because the database did not contain these outcome data. Further study is also needed to determine whether other factors such as ethnicity or complications are prognostic predictors of 5-year recurrence. Nevertheless, the results can still be considered valid given the robustness and statistical significance of the results.

Conclusions

Due to high recurrence rates in patients with HCC, accurate prediction of disease recurrence is of major importance for effectively tailoring adjuvant treatment and follow-up regimes for individual patients. The proposed deep-learning DNN model appears promising for use in prediction and stratification of 5-year recurrence risk after HCC resection based on clinical data. Predictors of postoperative 5-year recurrence can be discussed when educating HCC candidates in the expected course of recovery and health outcomes. Although the practical applicability of database studies such as this

Table 7. Recurrence risk factors after hepatocellular carcinoma resection: factors reported in selected studies

Authors (country)	No. of subjects	Measures	Findings
Shi et al., 2021 [the present study]	725 hepatocellular carcinoma (HCC) patients who had undergone surgical resection	Beck Depression Scale (BDS) score and Beck Anxiety Scale (BAS) score	The most important potential predictor in 5-year recurrence after HCC resection was surgeon volume followed by, in order of importance, hospital volume, preoperative BDS score, preoperative BAS score, co-residence with family, tumor stage, and tumor size.
Chiu et al., 2015 (Taiwan) [29]	23,107 patients who received HCC surgery		Overall 90-day mortality and 5-year mortality were significantly associated with treatment at a high-volume hospital (hazards ratio, HR = 0.79, 95% confidence interval (CI) 0.70-0.89; HR = 0.91, 95% CI 0.87-0.97, respectively) and treatment by a low-volume (HR = 0.86, 95% CI 0.82-0.90; HR = 0.84, 95% CI 0.80-0.88, respectively) (all $P < 0.001$).
Oh & Son, 2021 [30]	6 studies from 35 relevant articles were selected for this study (26,329 cancer patients)	Hospital anxiety and depression scale (HADS), Hamilton rating scale for anxiety (HAMA), cognitive based stress management (CBSM)	1. A total of 31% of cancer patients show psychological stress-related clinical symptoms such as depression, anxiety, and psychotic disorders. 2. A meta-analysis proved the positive effects of these psychotherapies on the reduction in cancer recurrence risk in cancer patients (HR = 0.52, 95% CI 0.33-0.84).
Westmaas et al., 2020 (U.S.A) [31]	Participants were 1,255 registered users of The American Cancer Society Cancer Survivors Network	The Cancer Survivors Network (CSN)	1. Overall forty-five percent of survivors also reported simultaneously being a caregiver to someone else with cancer. 2. It indicated that after controlling for medical and sociodemographic variables, online ($\beta = 0.31$, $P = 0.001$) social support were associated with higher ratings of wellbeing.
Tsai et al., 2021 (Taiwan) [32]	2,137 patients with HCC surgery		Overall number of tumors ($P = 0.001$) and tumor stage ($P = 0.007$) were independent prognostic factors for disease-free survival. Additionally, tumor stage ($P < 0.001$) was independent risk factors for poor overall survival.
Kim et al., 2021 (South Korea) [33]	153 patients combined hepatocellular carcinoma and holangiocarcinoma (cHCC-CCA)		1. The 1-, 3-, and 5-year tumor recurrence rates were 31.8% and 49.8%, and 59.0%, respectively. 2. The 8th American Joint Committee on Cancer tumor stage is an significant factor of the tumor recurrence ($P < 0.001$).

have been convincingly demonstrated in the literature, future studies can expand the range of clinical variables included in the analysis, which could obtain additional novel results and potentially improve prediction accuracy. Further experiments using the proposed DNN model would clarify its potential uses for developing, promoting, and improving health policies for treating HCC patients after surgery.

Acknowledgements

This work was financially supported by grants from Chi Mei Medical Center, Liouying (CLFHR 10931) and from Ministry of Science and Technology (MOST 108-2410-H-037-006-SS3) in Taiwan. The funding source of the study had no role in the study design, data collection, analysis, or interpretation or writing of the report.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Hao-Hsien Lee, Department of Surgery, Chi Mei Medical Center, No. 201, Taikang, Taikang Vil., Liouying District, Tainan 73657, Taiwan. Tel: +886-6-622-6999#72008; E-mail: Hao_Hsien@hotmail.com

References

[1] Shin SW, Ahn KS, Kim SW, Kim TS, Kim YH and Kang KJ. Liver resection versus local ablation therapies for hepatocellular carcinoma within the milan criteria: a systematic review and meta-analysis. Ann Surg 2021; 273: 656-666.

[2] Xu XF, Xing H, Han J, Li ZL, Lau WY, Zhou YH, Gu WM, Wang H, Chen TH, Zeng YY, Li C, Wu MC, Shen F and Yang T. Risk factors, patterns, and outcomes of late recurrence after liver resection for hepatocellular carcinoma: a multicenter study from China. JAMA Surg 2019; 154: 209-217.

[3] Gelli M, Sebagh M, Porcher R, Romanelli E, Vibert E, Cunha AS, Castaing D, Rosmorduc O, Samuel D, Adam R and Cherqui D. Liver resection for early hepatocellular carcinoma: preoperative predictors of non transplantable recurrence and implications for treatment allocation. Ann Surg 2020; 272: 820-826.

[4] Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, Topol EJ, Ioannidis JPA, Collins GS and Maruthappu M. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020; 368: m689.

[5] Jin J, Yao Z, Zhang T, Zeng J, Wu L, Wu M, Wang J, Wang Y, Yu J and Zheng R. Deep learning radiomics model accurately predicts hepatocellular carcinoma occurrence in chronic hepatitis B patients: a five-year follow-up. Am J Cancer Res 2021; 11: 576-589.

[6] Tang S, Ou J, Liu J, Wu YP, Wu CQ, Chen TW, Zhang XM, Li R, Tang MJ, Yang LQ, Tan BG, Lu FL and Hu J. Application of contrast-enhanced CT radiomics in prediction of early recurrence of locally advanced oesophageal squamous cell carcinoma after trimodal therapy. Cancer Imaging 2021; 21: 38.

[7] Nam JY, Lee JH, Bae J, Chang Y, Cho Y, Sinn DH, Kim BH, Kim SH, Yi NJ, Lee KW, Kim JM, Park JW, Kim YJ, Yoon JH, Joh JW and Suh KS. Novel model to predict HCC recurrence after liver transplantation obtained using deep learning: a multicenter study. Cancers (Basel) 2020; 12: 2791.

[8] Ji GW, Zhu FP, Xu Q, Wang K, Wu MY, Tang WW, Li XC and Wang XH. Machine-learning analysis of contrast-enhanced CT radiomics predicts recurrence of hepatocellular carcinoma after resection: a multi-institutional study. EBioMedicine 2019; 50: 156-165.

[9] Liu IT, Yen CS, Wang WL, Tsai HW, Chu CY, Chang MY, Hou YF and Yen CJ. Predict early recurrence of resectable hepatocellular carcinoma using multi-dimensional artificial intelligence analysis of liver fibrosis. Cancers (Basel) 2021; 13: 5323.

[10] Mai RY, Zeng J, Meng WD, Lu HZ, Liang R, Lin Y, Wu GB, Li LQ, Ma L, Ye JZ and Bai T. Artificial neural network model to predict post-hepatectomy early recurrence of hepatocellular carcinoma without macroscopic vascular invasion. BMC Cancer 2021; 21: 283.

[11] Song D, Wang Y, Wang W, Wang Y, Cai J, Zhu K, Lv M, Gao Q, Zhou J, Fan J, Rao S, Wang M and Wang X. Using deep learning to predict microvascular invasion in hepatocellular carcinoma based on dynamic contrast-enhanced MRI combined with clinical parameters. J Cancer Res Clin Oncol 2021; 147: 3757-3767.

[12] Beck AT and Steer RA. Beck anxiety inventory manual. San Antonio, TX: Psychological Corporation; 1993.

[13] Beck AT, Steer RA and Brown G. Beck depression inventory (BDI-II) manual. San Antonio, TX: Psychological Corporation, Harcourt Brace; 1996.

[14] Lee HH, Chiu CC, Lin JJ, Wang JJ, Lee KT, Sun DP and Shi HY. Impact of preoperative anxiety and depression on quality of life before and after resection of hepatocellular carcinoma. J Affect Disord 2019; 246: 361-367.

[15] LeCun Y, Bengio Y and Hinton G. Deep learning. Nature 2015; 521: 436-444.
[16] Yu Y, Si X, Hu C and Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput 2019; 31: 1235-1270.
[17] Altmann A, Toloşi L, Sander O and Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics 2010; 26: 1340-1347.
[18] Chang CM, Yin WY, Wei CK, Lee CH and Lee CC. The combined effects of hospital and surgeon volume on short-term survival after hepatic resection in a population-based study. PLoS One 2014; 9: e86444.
[19] Franchi E, Donadon M and Torzilli G. Effects of volume on outcome in hepatobiliary surgery: a review with guidelines proposal. Glob Health Med 2020; 2: 292-297.
[20] Lu CC, Chiu CC, Wang JJ, Chiu YH and Shi HY. Volume-outcome associations after major hepatectomy for hepatocellular carcinoma: a nationwide Taiwan study. J Gastrointest Surg 2014; 18: 1138-1145.
[21] Wang X, Wang N, Zhong L, Wang S, Zheng Y, Yang B, Zhang J, Lin Y and Wang Z. Prognostic value of depression and anxiety on breast cancer recurrence and mortality: a systematic review and meta-analysis of 282,203 patients. Mol Psychiatry 2020; 25: 3186-3197.
[22] Luo X, Li W, Yang Y, Humphris G, Zeng L, Zhang Z, Garg S, Zhang B and Sun H. High fear of cancer recurrence in chinese newly diagnosed cancer patients. Front Psychol 2020; 11: 1287.
[23] Chen SJ, Chang CH, Chen HC and Liu CY. Association between depressive disorders and risk of breast cancer recurrence after curative surgery. Medicine (Baltimore) 2016; 95: e4547.
[24] Su JA, Yeh DC, Chang CC, Lin TC, Lai CH, Hu PY, Ho YF, Chen VC, Wang TN and Gossop M. Depression and family support in breast cancer patients. Neuropsychiatr Dis Treat 2017; 13: 2389-2396.
[25] Kamen C, Cosgrove V, McKellar J, Cronkite R and Moos R. Family support and depressive symptoms: a 23-year follow-up. J Clin Psychol 2011; 67: 215-223.
[26] Given BA, Given CW and Kozachik S. Family support in advanced cancer. CA Cancer J Clin 2001; 51: 213-231.
[27] Zhang Y, Chen SW, Liu LL, Yang X, Cai SH and Yun JP. A model combining TNM stage and tumor size shows utility in predicting recurrence among patients with hepatocellular carcinoma after resection. Cancer Manag Res 2018; 10: 3707-3715.
[28] Kumar AM, Fredman ET, Coppa C, El-Gazzaz G, Aucejo FN and Abdel-Wahab M. Patterns of cancer recurrence in localized resected hepatocellular carcinoma. Hepatobiliary Pancreat Dis Int 2015; 14: 269-275.
[29] Chiu CC, Wang JJ, Chen YS, Chen JJ, Tsai TC, Lai CC, Sun DP and Shi HY. Trends and predictors of outcomes after surgery for hepatocellular carcinoma: a nationwide population-based study in Taiwan. Eur J Surg Oncol 2015; 41: 1170-1178.
[30] Oh HM and Son CG. The risk of psychological stress on cancer recurrence: a systematic review. Cancers (Basel) 2021; 13: 5816.
[31] Westmaas JL, Fallon E, McDonald BR, Driscoll D, Richardson K, Portier K and Smith T. Investigating relationships among cancer survivors' engagement in an online support community, social support perceptions, well-being, and moderating effects of existing (offline) social support. Support Care Cancer 2020; 28: 3791-3799.
[32] Tsai YC, Sou FM, Liu YW, Wu YJ, Yong CC, Chen DW, Huang PY, Cho WR, Chuang CH, Hsiao CC, Hu TH and Tsai MC. Preoperative ALBI grade predicts the outcomes in non-B non-C HCC patients undergoing primary curative resection. BMC Gastroenterol 2021; 21: 386.
[33] Kim M, Hwang S, Ahn CS, Kim KH, Moon DB, Song GW, Jung DH and Hong SM. Post-resection prognosis of combined hepatocellular carcinoma-cholangiocarcinoma cannot be predicted by the 2019 World Health Organization classification. Asian J Surg 2021; 44: 1389-1395.