The diagnostic nomogram of platelet-based score models for hepatic alveolar echinococcosis and atypical liver cancer

Qiancheng Du1,6, Yanyan Wang2,6, Shihao Guan3,6, Chenliang Hu4,6, Mengxuan Li6, Ling Zhou1, Mengzhao Zhang6, Yichong Chen5, Xuepeng Mei5, Jian Sun1,7* & Ying Zhou4,7*

Hepatic alveolar echinococcosis (HAE) and liver cancer had similarities in imaging results, clinical characteristics, and so on. And it is difficult for clinicians to distinguish them before operation. The aim of our study was to build a differential diagnosis nomogram based on platelet (PLT) score model and use internal validation to check the model. The predicting model was constructed by the retrospective database that included in 153 patients with HAE (66 cases) or liver cancer (87 cases), and all cases was confirmed by clinicopathology and collected from November 2011 to December 2018. Lasso regression analysis model was used to construct data dimensionality reduction, elements selection, and building prediction model based on the 9 PLT-based scores. A multi-factor regression analysis was performed to construct a simplified prediction model, and we added the selected PLT-based scores and relevant clinicopathologic features into the nomogram. Identification capability, calibration, and clinical serviceability of the simplified model were evaluated by the Harrell’s concordance index (C-index), calibration plot, receiver operating characteristic curve (ROC), and decision curve. An internal validation was also evaluated by the bootstrap resampling. The simplified model, including in 4 selected factors, was significantly associated with differential diagnosis of HAE and liver cancer. Predictors of the simplified diagnosis nomogram consisted of the API index, the FIB-4 index, fibro-quotent (FibroQ), and fibrosis index constructed by King's College Hospital (King's score). The model presented a perfect identification capability, with a high C-index of 0.929 (0.919 through internal validation), and good calibration. The area under the curve (AUC) values of this simplified prediction nomogram was 0.929, and the result of ROC indicated that this nomogram had a good predictive value. Decision curve analysis showed that our differential diagnosis nomogram had clinically identification capability. In conclusion, the differential diagnosis nomogram could be feasibly performed to verify the preoperative individualized diagnosis of HAE and liver cancer.

Hepatic alveolar echinococcosis (HAE), also known as hepatic malignant parasitic disease, is an ancient zoonotic parasitic disease1. In recent years, HAE has become a worldwide epidemic disease that seriously endangers the world’s public health and economic development with the development of tourism, the flow of population and the rapid increase of domestic dogs2. HAE proliferates by the means of budding or infiltration, which could produce new vesicles that infiltrated into the deep tissue. HAE could not only directly infiltrate the adjacent tissue, but also could be moved to the peritoneum and distant organs through lymphatic channels and blood vessels3. The diagnosis of HAE mainly depends on epidemiological evidence, clinical characteristics, serology and immunological

1Department of General Surgery, Shanghai Fourth People’s Hospital Affiliated to Tongji University School of Medicine, Shanghai, 200081, China. 2Department of Hematology, Fuyang Hospital of Anhui Medical University, Fuyang, 236000, China. 3Department of Plastic surgery, the Qingpu Branch of Zhongshan Hospital Affiliated to Fudan University, Shanghai, 201700, China. 4Department of Hepatopancreatobiliary Surgery, the Affiliated Hospital of Qinghai University and Qinghai Province Key Laboratory of Hydatid Disease Research, Qinghai University, Xining, 81000, China. 5The Graduate School, Qinghai University, Xining, 81000, China. 6These authors contributed equally: Qiancheng Du, Yanyan Wang, Shihao Guan and Chenliang Hu. 7These authors jointly supervised this work: Jian Sun and Ying Zhou. *email: 13311856338@163.com; 17697227511@163.com
Retrieved from Scientific Reports, DOI: 10.1038/s41598-019-55563-3
Identification capability of differential diagnosis nomogram. The calibration curve of the prediction nomogram for the differential diagnosis of HAE and liver cancer presents a good agreement in our research (Fig. 5A). The Harrell’s concordance index (C-index) for our prediction nomogram is 0.929 (0.886–0.972) for our cohort. The C-index of our nomogram is validated to be 0.919 through bootstrapping validation, and the result also demonstrated a good identification capability. The ROC curve of the differential diagnosis nomogram in the cohort is showed in Fig. 5B. The AUC values of this prediction nomogram is 0.929. The result of ROC indicates that this nomogram have a good predictive value.

The decision curve for the differential diagnosis nomogram is showed in Fig. 5C. The decision curve shows that if the threshold probability of a patient or doctor is >6%, using the nomogram to diagnose HAE or liver cancer could acquire much more benefit. And it is obvious for net benefit of the differential diagnosis nomogram within this range of threshold probability.

Discussion

It was difficult to distinguish HAE from liver cancer because HAE had no specific clinical features, and imaging examination also lacked characteristic finding. The imaging findings of HAE and liver cancer are similar, such as a large space-occupying lesion in the liver, irregular areas of reduced density, etc.14–16. Sometimes, liver cancer is difficult to distinguish from HAE in imaging, especially for those patients with low levels of AFP. Atypical liver cancer presents cystic lesion, no pseudocapsule, no dynamic scanning enhancement, and no portal vein invasion or tumor thrombus formation in imaging finds. The gold standard that diagnosed liver lesions is histopathological.

Table 1. Baseline Characteristics of HAE and liver cancer. Abbreviation: HAE, hepatic alveolar echinococcosis; INR, international normalized ratio; PT, Prothrombin time; ALT, Alanineaminotransferase; AST, Aspartateaminotransferase; PLT, Platelet count; API, Age/platelet count index; CDS, Cirrhosis discriminant score; FIB-4, Fibrosis index based on the four factors; FibroQ, Fibro-quotient; GUCI, Goteburg University Cirrhosis Index; King's score, Fibrosis index based on the four factors; AAR, Aspartate aminotransferase/Alanine aminotransferase ratio; AARP, AAR-platelet count score.
Figure 1. Receiver operating characteristic analysis using PLT, GUCI, FibroQ, FIB-4, King's score, CDS, API, and their combination in HAE group (n = 66) and liver cancer group (n = 87).

Figure 2. PLT-based score models selection using the LASSO binary logistic regression model. (A) The Optimum parameter (lambda) selection in the LASSO model performed fivefold cross-validation through minimum criteria. The partial likelihood deviance (binomial deviance) curve was presented versus log (lambda). Dotted vertical lines were showed at the optimum values by performing the minimum criteria and the 1-SE of the minimum criteria (the 1-SE criteria). (B) The LASSO coefficient profiles of the 9 features. The coefficient profile plot was evaluated against the log (lambda) sequence. Vertical line was shown at the value selected using cross-validation, where the optimum lambda gave rise to four features with nonzero coefficients.

Figure 3. The results of the logistic regression analysis among API, FIB-4, FibroQ, and King's score that selected by the LASSO regression model were presented in forest plot.
Figure 4. Developed differential diagnosis nomogram for distinguishing HAE and liver cancer. To utilize the nomogram, an individual patient's value was presented on each variable axis, and a vertical line was drawn upward to find the number of points received for each variable value. The sum of the variable values was presented on the total point axis, and a vertical line was also drawn downward to the differential diagnosis axes to seek the probability of liver cancer.
accuracy of diagnosis on preoperative peripheral blood tumor marker, such as carcinoembryonic antigen (CEA) and peripheral inflammatory factors, for lymph node metastasis of colorectal cancer was less than 70%, much lower than the C-index of PLT-based score models we constructed24. Therefore, the noninvasive, simplified pre- diagnosis nomogram that incorporated the routine laboratory tests we already get for available, which could regard as a more simplified model for distinguishing HAE from liver cancer.

Both doctors and patients could also perform an individualized diagnosis model for predicting the probability of HAE and liver cancer with the easy feasible scoring system, which is in line with the current concept of precision medicine25. However, the clinical outcomes, the particular level of identification or degree of calibration, could not be got by the prediction capability, identification capability and calibration of nomogram26,27. The decision curve of our nomogram presents that if the threshold probability of an individual is more than 6%, using the nomogram to diagnose HAE or liver cancer could acquire much more benefit. Within this range, net benefit was comparable, with several overlaps, on the basis of the differential diagnosis nomogram.

Conclusions
This research presents a differential diagnosis nomogram that incorporates the PLT-based score models, which could be easily implemented to promote the pre-therapy individualized diagnosis of HAE and liver cancer.

Methods
This study was approved by the Institutional Research Ethics Board of Affiliated Hospital of Qinghai University, and all methods were performed in accordance with the Declaration of Helsinki (P-SL-2018005). All patients signed the written informed consent before surgery. This study did not involve human or animal tissue or blood samples.

Patients. We retrospectively collected the demographic, clinical characteristics, and peripheral blood data of 153 patients in the Department of Hepatopancreatobiliary Surgery at the Affiliated Hospital of Qinghai University (Qinghai, China) and the Department of General Surgery at the Shanghai Fourth People’s Hospital (Shanghai, China) between November 2011 and December 2018. The 153 patients was divided into two groups: (i) HAE group (66) recruited from the Affiliated Hospital of Qinghai University and (ii) liver cancer group (87) recruited from the two hospitals. All peripheral blood parameters are derived from blood draws taken within 7 days before all kinds of therapies. All patients, including HAE and liver cancer, have been confirmed by surgery or ultrasound guided fine needle aspiration biopsy. Atypical liver cancer is defined that 2 doctors (at least the associate chief physician) are unable to definitively diagnose liver cancer or HAE according to imaging characteristics and peripheral blood tumor markers before surgery. The CT features of atypical liver cancer are cystic lesion, no pseudocapsule, no dynamic scanning enhancement, and no portal vein invasion or tumor thrombus formation. Patients who have one of the following would be excluded: (i) combined with imaging characteristics and peripheral blood tumor markers, the diagnosis could be confirmed; (ii) patients with hematological system diseases; (iii) patients have received the blood transfusion within the previous 6 months; (iv) patients with incomplete clinical data; (v) patients comorbid with non-alcoholic steatohepatitis (NASH), long-term alcohol, long-term oral drugs, etc. The coding system is performed to assure the anonymity of all patients enrolled into the study.

Platelet-based score models. Electronic medical records are used to obtain relevant information, including age, sex, levels of alanine aminotransferase (ALT), aspartate aminotransferase (AST), PLT, international
The LASSO, particularly suits to the reduction in high dimensional data, is performed to choose the optimum predictive factors in the differential diagnosis of HAE and liver cancer. Factors with nonzero coefficients in the LASSO regression model should be chosen. Multivariate analysis is performed to construct a predicting model by combining the factors selected in the LASSO regression model. The result is presented in a figure using the stata15.0 (https://www.stata.com/). Figure is characterized by odds ratio (OR) and 95%CI. Variables chose by LASSO regression model are included in the model. All possible diagnostic factors are performed to construct a simplified model for the differential diagnosis of HAE and liver cancer. A simplified nomogram is showed based on the results of LASSO regression model. The Harrell's concordance index (C-index) is performed to quantify the discrimination performance of the simplified nomogram. In general, the value of C-index greater than 0.75 is regarded as a relatively good discrimination. The calibration curves are plotted to evaluate the predictive ability of the nomogram related with platelet-based models. The bootstrapping method is performed to reduce estimate bias. The predictive nomogram of platelet-based models is conducted to bootstrapping validation (10000 resample reseamps) to count a relatively corrected C-index. The decision curve is presented to evaluate the net benefit of different threshold probabilities for individuals to confirm clinical usefulness of the simplified nomogram. All of these statistical methods are performed by R software version 3.6.0 (http://www.r-project.org) with R packages "rms", "glmnet", "Hmisc", "ROCR", "rmda", "caret", and "foreign".

### Table 2. Scoring of platelet-based models

| Index          | Formulas                                                                 |
|----------------|--------------------------------------------------------------------------|
| Pohl score     | 1: AAR > 1 and PLT < 150 × 10^9/L or else, the score = 0                |
| AARP           | 1: AAR > 1 or PLT < 150 × 10^9/L or else, the score = 0                |
| API            | Age (years): < 30 = 0; 30–39 = 1; 40–49 = 2; 50–59 = 3; 60–69 = 4; ≥ 70 = 5; PLT: ≥ 225 = 0; 200–224 = 1; 175–199 = 2; 150–174 = 3; 125–149 = 4; < 125 = 5. API is the sum of age and platelet scores and therefore varied from 0–10 |
| CDS            | PLT: ≥ 340 = 0; 280–339 = 1; 220–279 = 2; 160–219 = 3; 100–159 = 4; 40–99 = 5; < 40 = 6. ALT/AST ratio: ≥ 1.7 = 0; 1.2–1.7 = 1; 0.6–1.1 = 2; < 0.6 = 3. INR: < 1.1 = 0; 1.1–1.4 = 1; 1.4–2 = 2. CDS is the sum of the above |
| FIB-4          | Age (years) × AST (U/L) / [PLT (109/L) × ALT (U/L)] × 25 |
| FibroQ         | 10 × (Age × ALT/ INR × AL/PLT)                                         |
| GUCI           | AST × INR × 100/PLT (10^9/L)                                           |
| King's score   | Age × AT × INR/PLT (10^9/L)                                             |

**Statistical analysis.** Statistical analysis of the numerical variables among the two groups is performed using unpaired Student’s t-test for parametric data or Mann-Whitney rank sum test for nonparametric data. P-value < 0.05 is considered that the difference is statistically significant. The specificity and sensitivity are evaluated by the numerical integration of ROC. And the optimal cut-off point of each variable, the sum of specificity and sensitivity is the highest cumulative value, is obtained by the ROC. Figures are plotted by the GraphPad Prism 5.01 (https://www.graphpad.com/). A bilateral P value < 0.05 is considered statistically significant. All above analyses are performed using SPSS 23.0 software (IBM Corporation, 2015, USA).

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 31 July 2019; Accepted: 29 November 2019; Published: 18 December 2019

**References**

1. Carmona, D. & Cardona, G. A. Echinococcosis in wild carnivorous species: epidemiology, genotypic diversity, and implications for veterinary public health. *Vet Parasitol* **202**, 69–94 (2014).
2. Farrokhi, D., Zandi, B., Pezeshki Rad, M. & Tavakoli, M. Hepatic Alveolar Echinococcosis. *Arch Iran Med* **18**, 199–202 (2015).
3. Lötsch, F. et al. FDG-PET/MRI imaging for the management of alveolar echinococcosis: initial clinical experience at a reference centre in Austria. *Trop Med Int Health* **24**, 663–670 (2019).
4. Filipou, D., Tsellepis, D., Filippou, G. & Papadopoulos, V. Advances in liver echinococcosis: diagnosis and treatment. *Clin Gastroenterol Hepatol* **5**, 152–159 (2007).
5. Du, Q. C., Zhou, Y. & Fan, H. N. Treatment of hepatic alveolar echinococcosis infringing the inferior vena cava. *Rev Soc Bras Med Trop* **52**, e20180342 (2019).
RETRACTED ARTICLE

Zhou, Mengzhao Zhang, Xuepeng Mei and Yichong Chen: data collection. manuscript writing. Jian Sun and Ying Zhou: manuscript review. Mengxuan Li: data collection and analysis. Ling Qiancheng Du, Yanyan Wang, Shihao Guan and Chenliang Hu: study design, data collection, analysis, and Surgery, the Affiliated Hospital of Qinghai University. Research reported in this publication was supported by the Departments of Hepatopancreatobiliary

The authors would like to thank all colleagues for data collection from the Departments of Hepatopancreatobiliary Surgery, the Affiliated Hospital of Qinghai University. Research reported in this publication was supported by grant number SY-XKZT-2019–2004 from the promoting subject development program of Shanghai fourth People's Hospital.

Author contributions
Qiancheng Du, Yanyan Wang, Shihao Guan and Chenliang Hu: study design, data collection, analysis, and manuscript writing. Jian Sun and Ying Zhou: manuscript review. Mengxuan Li: data collection and analysis. Ling Zhou, Mengzhao Zhang, Xuepeng Mei and Yichong Chen: data collection.
Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to J.S. or Y.Z.
Reprints and permissions information is available at www.nature.com/reprints.
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019