Development and validation of a radiomics-based method for macrovascular invasion prediction in hepatocellular carcinoma with prognostic implication

Jingwei Wei[a, b, c, #], Sirui Fu[d, #], Shuaitong Zhang[a, b, c], Jie Zhang[e], Dongsheng Gu[a, b, c], Xiaoqun Li[f], Xudong Chen[g], Xiaofeng He[h], Jianfeng Yan[i], Ligong Lu[d, *], Jie Tian[a, b, c, *]

[a]Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b]Beijing Key Laboratory of Molecular Imaging, Beijing 100190, China
[c]University of Chinese Academy of Sciences, Beijing 100049, China
[d]Center of Intervention Radiology, Center of Precise Medicine, Zhuhai People's Hospital, Zhuhai, China
[e]Department of Radiology, Center of Precise Medicine, Zhuhai People's Hospital Zhuhai, China
[f]Department of Interventional Treatment, Zhongshan City People's Hospital, Zhongshan, China
[g]Department of Radiology, Shenzhen People's Hospital, Shenzhen, China
[h]Interventional Diagnosis and Treatment Department, Nanfang Hospital, Southern Medical University, Guangzhou, China
[i]Department of Radiology, Yangjiang People's hospital, Yangjiang, China

ABSTRACT

In hepatocellular carcinoma (HCC), more than one third of patients were accompanied by macrovascular invasion (MaVI) during diagnosis and treatment. HCCs with MaVI presented with aggressive tumor behavior and poor survival. Early identification of HCCs at high risk of MaVI would promote adequate preoperative treatment strategy making, so as to prolong the patient survival. Thus, we aimed to develop a computed tomography (CT)-based radiomics model to preoperatively predict MaVI status in HCC, meanwhile explore the prognostic prediction power of the radiomics model.

A cohort of 452 patients diagnosed with HCC was collected from 5 hospitals in China with complete CT images, clinical data, and follow-ups. 15 out of 708 radiomic features were selected for MaVI prediction using LASSO regression modeling. A radiomics signature was constructed by support vector machine based on the 15 selected features. To evaluate the prognostic power of the signature, Kaplan-Meier curves with log-rank test were plotted on MaVI occurrence time (MOT), progression free survival (PFS) and overall survival (OS).

The radiomics signature showed satisfactory performance on MaVI prediction with area under curves of 0.885 and 0.770

# Equal contributors: Jingwei Wei and Sirui Fu
* Equal corresponding authors: Jie Tian, Email: tian@ieee.org; Ligong Lu, Email: llg0902@sina.com

on the training and external validation cohorts, respectively. Patients could successfully be divided into high- and low-risk groups on MOT and PFS with p-value of 0.0017 and 0.0013, respectively. Regarding to OS, the Kaplan-Meier curve did not present with significant difference which may be caused by non-uniform following treatments after disease progression.

To conclude, the proposed radiomics model could facilitate MaVI prediction along with prognostic implication in HCC management.

Keywords: hepatocellular carcinoma; macrovascular invasion; prognosis; computed tomography; radiomics; prediction

1. DESCRIPTION OF PURPOSE

Hepatocellular carcinoma (HCC) is the most common liver primary tumor and leads to over 500,000 deaths worldwide each year[1-3]. Macrovascular invasion (MaVI) occurs in up to 50% HCC patients during diagnosis and treatment[4-6]. The presence of MaVI indicates severe disease progression and usually corresponds to postoperative recurrence[7-9]. Currently, the diagnosis of MaVI is through radiological examination after existence, which limits prevention measures and adequate preoperative decision making. Accurate and robust method for MaVI prediction is still in need.

Previous studies have reported clinical factors for MaVI prediction, including tumor size, nodule number, and alpha-fetoprotein level, etc[6, 10]. Although these clinical factors have close correlation with MaVI, the predictive accuracy remains low. Considering the advantage of radiological imaging in HCC management, method based on medical imaging is worth exploring[11, 12]. Radiomics provides a possible solution to this problem. Radiomics depicts tumor heterogeneity through extracting thousands of quantitative imaging features. Machine learning methods are used to construct the final predictive radiomics model by integrating efficient features though big data mining[13-17].

Thus, in this multicenter study, we aimed to develop a preoperative prediction model for MaVI using radiomics pipeline. Computed tomography (CT) – based imaging analysis was implemented to construct the radiomics signature for MaVI occurrence discrimination. Furthermore, the prognostic prediction power of the radiomics signature was validated on MaVI occurrence time (MOT), progression free survival (PFS), and overall survival (OS).

2. METHODS

2.1 Patients and data acquisition

This retrospective study was approved by the institutional review board. We reviewed a total of 452 patients diagnosed with HCC from 5 independent hospitals in China from April 2007 to November 2016. The cooperative hospitals are accordingly: Yangjiang People's Hospital, Zhongshan City People's Hospital, Zhuhai People's Hospital, Shenzhen People's Hospital, and Nanfang Hospital. CT imaging data and demographic/clinical data were acquired from picture archiving and communication systems and electronic medical reports, respectively. MaVI was confirmed by CT or magnetic resonance imaging. Follow-ups included PFS, OS, and MOT.

2.2 Datasets division

The enrolled cohort was divided into three groups: 1) No MaVI occurrence from initial diagnosis to the end of follow-up (n = 287); 2) MaVI occurrence at initial diagnosis (n = 114); 3) No MaVI occurrence at initial diagnosis, but occurred during the follow-up (n = 47), which were accordingly named as Non-MaVI, MaVI, and MaVI occurrence groups. To better select correlated imaging features with maximum discriminative power, non-MaVI and MaVI groups were chosen

for radiomics signature construction. The training dataset included 172 non-MaVI patients and 82 MaVI patients. The validation dataset included 115 non-MaVI patients and 32 MaVI patients. Prognosis related analysis were performed on MaVI occurrence group regarding to PFS, OS, and MOT.

2.3 Development and validation of the radiomics-based model

A set of 708 radiomic features were extracted from CT imaging on portal phase. The extracted features could be divided into 5 types including first-order, textural, Gabor wavelet, semantic, and peel-off features[18, 19]. Semantic features were designed based on radiologists' experience, which were accordingly tumor border clearness, necrosis area, circularity, number of nodule heaves on the border, difference between center and cavity, and mosaic area. Peel-off features were 10 statistical features extracted from peel-off layers which contain voxels from outside to insider tumor. Firstly, redundant features were removed by setting correlation coefficients larger than 0.75[20]. Then, lasso-logistic regression modeling was employed to select efficient features from the feature pool[21]. The radiomics signature was finally constructed by integrating the selected features using linear support vector machine. Receiver operating characteristic (ROC) curve was used to show the classification ability of the radiomics signature[22]. Area under the curve (AUC), accuracy, specificity, and sensitivity were utilized to assess the predictive performance. To evaluate the prognostic power of the signature, we drew Kaplan-Meier curves regarding to PFS, OS, and MOT in the MaVI occurrence group[23]. Log-rank test manifested whether there existed significant differences between the high-risk and low-risk subgroups.

3. RESULTS

The radiomics signature was constructed by 15 selected radiomic features. It manifested satisfactory predictive performance with AUCs of 0.8849 and 0.7704 in the training and validation datasets, respectively. The ROC curves are shown in Figure 1. The predictive accuracy was 75.98% in the training dataset and 73.47% in the validation datasets. Other detailed predictive indicators are shown in Table 1.

Table 1. Predictive indicators of the radiomics signature

Predictive indicator	Training dataset (n = 254)	Validation dataset (n = 147)
AUC	0.8849	0.7704
Accuracy	75.98%	73.47%
Specificity	0.6919	0.7304
Sensitivity	0.9024	0.7500

For prognostic analysis, the radiomics signature could successfully divide patients into high-risk and low-risk subgroups regarding to MOT and PFS with p-value of 0.0017 and 0.00134, respectively, in MaVI occurrence group. However, OS risk stratification did not show significant difference between the divided subgroups with p-value of 0.14405. The Kaplan-Meier curves of MOT, PFS, and OS in MaVI occurrence group are shown in Figure 2.

4. NEW OR BREAKTHROUGH WORK TO BE PRESENTED

The originality of this work lies in that we figured out MaVI correlated radiological features. Compared with previously reported clinical factors, the constructed radiomics signature had better performance on MaVI prediction with higher AUC. In addition, we revealed the prognostic power of the proposed signature on MOT, as well as PFS, which would facilitate identification of high-risk patients in HCC management, thus assisting personalized treatment decision making.

5. CONCLUSIONS

In HCC management, MaVI occurrence represents uncontrollable progression dilemma. In this multicenter study, we developed and validated a radiomics-based method to preoperatively predict MaVI. The proposed radiomics signature could successfully realize MaVI prediction with improved predictive performance compared with conventional clinical factors. The radiomics signature also presented with satisfactory prognostic implications on MOT and PFS, which could guide identification of high-risk patients who should be provided with more frequent follow-ups and earlier treatment intervention. The proposed radiomics method would beyond doubt provide reliable basis for HCC management in clinical settings.

6. ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China (2017YFA0205200, 2017YFC1308700, 2017YFC1308701, 2017YFC1309100, 2016YFC0103803, and 2017YFA0205200), National Natural Science Foundation of China (81227901, 81771924, 81501616, 81571785, 81771957, and 61671449), the Natural Science Foundation of Guangdong Province, China (No. 2016A030311055, 2016A030313770).

7. REFERENCES

1. Bosch FX, Ribes J, Díaz M, Cléries R, “Primary liver cancer: Worldwide incidence and trends”, Gastroenterology. 127:S5. (2014)
2. Parkin DM, Pisani P, Ferlay J, “Global cancer statistics”, Ca A Cancer Journal for Clinicians. 61:69. (2011)
3. Forner A, Llovet JM, Bruix J, “Hepatocellular carcinoma”, Lancet. 379:1245-1255. (2012)
4. Minagawa M, Makuuchi M, “Treatment of hepatocellular carcinoma accompanied by portal vein tumor thrombus”, World Journal of Gastroenterology. 12:7561. (2006)
5. Yeung YP, Lo CM, Liu CL, Wong BC, Fan ST, Wong J, “Natural History of Untreated Nonsurgical Hepatocellular Carcinoma”, American Journal of Gastroenterology. 100:1995-2004. (2005)
6. Sakata J, Shirai Y, Wakai T, Kaneko K, Nagahashi M, Hatakeyama K, “Preoperative predictors of vascular invasion in hepatocellular carcinoma”, European Journal of Surgical Oncology the Journal of the European Society of Surgical Oncology & the British Association of Surgical Oncology. 34:900. (2008)
7. Llovet JM, Schwartz M, Mazzaferro V, “Resection and Liver Transplantation for Hepatocellular Carcinoma”, Seminars in liver disease. 181-200. (2005)
8. Imamura H, Matsuyama Y, Tanaka E, et al, “Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy”, Journal of Hepatology. 38:200. (2003)
9. Jonas S, Bechstein WO, Steinmüller T, et al, “Vascular invasion and histopathologic grading determine outcome after liver transplantation for hepatocellular carcinoma in cirrhosis”, Hepatology. 33:1080. (2001)
10. Pawlik TM, Delman KA, Vauthey JN, et al, “Tumor size predicts vascular invasion and histologic grade: Implications

for selection of surgical treatment for hepatocellular carcinoma", Liver Transpl. 11:1086-1092. (2005)

11. Kurokawa T, Yamazaki S, Mitsuka Y, Moriguchi M, Sugitani M, Takayama T, "Prediction of vascular invasion in hepatocellular carcinoma by next-generation des-r-carboxy prothrombin", Br J Cancer. 114:53-58. (2016)

12. Ramos E, Lladó L, Serrano T, et al, "Utility of cell-cycle modulators to predict vascular invasion and recurrence after surgical treatment of hepatocellular carcinoma", Transplantation. 82:753. (2006)

13. Lambin P, Leijenaar R, Deist TM, et al, "Radiomics: the bridge between medical imaging and personalized medicine", Nature Reviews Clinical Oncology. 14:749. (2017)

14. Huang YQ, Liang CH, He L, et al, "Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer", Journal of Clinical Oncology Official Journal of the American Society of Clinical Oncology. 34:2157. (2016)

15. Segal E, Sirlin CB, Ooi C, et al, "Decoding global gene expression programs in liver cancer by noninvasive imaging", Nature Biotechnology. 25:675. (2007)

16. Bakr S, Echegaray S, Shah R, et al, "Noninvasive radiomics signature based on quantitative analysis of computed tomography images as a surrogate for microvascular invasion in hepatocellular carcinoma: a pilot study", J Med Imaging 4:041303. (2017)

17. Ying Z, Lan H, Huang Y, et al, "CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma", Abdominal Radiology. 42:1-10. (2017)

18. Aerts HJ, Velazquez ER, Leijenaar RT, et al, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach", Nature communications. 5. (2014)

19. Lambin P, Riosvelazquez E, Leijenaar R, et al, "Radiomics: Extracting more information from medical images using advanced feature analysis", European Journal of Cancer. 48:441. (2012)

20. Lin L, "A concordance correlation coefficient to evaluate reproducibility", Biometrics. 45:255-268. (1989)

21. Tibshirani R, "Regression shrinkage and selection via the lasso: a retrospective", Journal of the Royal Statistical Society. 58:267-288. (1996)

22. Zou K, O'Malley A, L, "Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models" Circulation. 115:654-657. (2007)

23. BradleyEfron, "Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve", Publications of the American Statistical Association. 83:414-425. (1988)

8. FIGURES

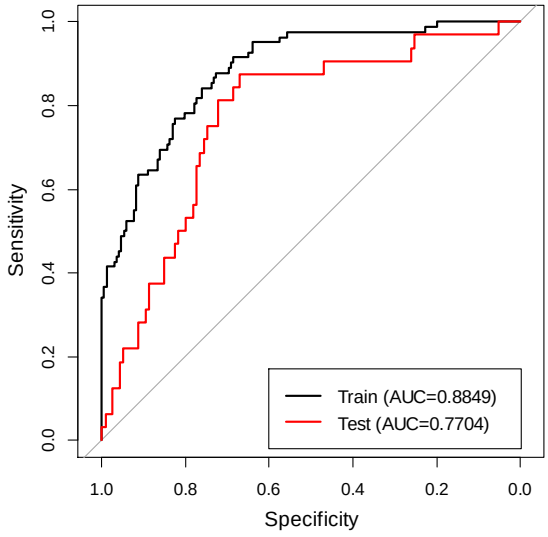

Figure 1. ROC curves of the radiomics signature. The black line represents for ROC curve in the training dataset. The red line represents for ROC curve in the validation dataset.

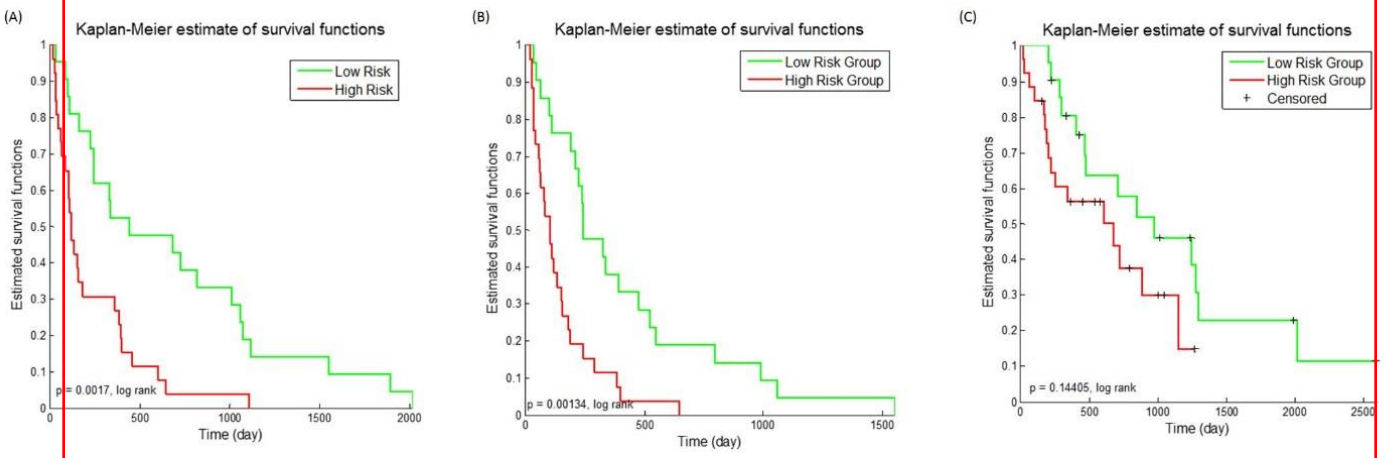

Figure 2. Kaplan-Meier curves in MaVI occurrence group. (A) Kaplan-Meier curve on MOT; (B) Kaplan-Meier curve on PFS; (C) Kaplan-Meier curve on OS. P-value less than 0.05 represents for significant difference.