ROUND1

Dear editor and reviewers:

Thank you very much for your comments.

Reply to the comments on Manuscript ID: 74196: Radiomic Analysis Based on Multi-phase MRI to Predict Preoperatively Microvascular Invasion in Hepatocellular Carcinoma. Yueming Li, Yuemin Zhu, Lanmei Gao, et al.

I have revised our manuscript according to all comments of the reviews. According with your advice, we amended the relevant part in manuscript. We resubmit the revised manuscript **polished by English professionals**, a document answering every question point by point from the reviewers, a Figures.pptx File, a Tables.docx File, a document of supplementary materials and English service letter. A revised manuscript with the correction sections red marked was used for easy check/editing purpose. Should you have any questions, please contact us without hesitate. Here below is our description on revision according to the reviewers' comments.

Reviewer #1:

Scientific Quality: Grade C (Good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Major revision

1) General Comments: In this manuscript, the authors aimed to show the higher accuracy of an algorithm using most discriminant factors (MDF), which are created by image texture analysis using MaZda software as 101 features and are statistically refined into 30 features, in prediction of microvascular invasion (MVI) comparing with a regular image diagnosis by radiologists. Basically, this manuscript is consisting of statistical analyses and MRI images. The statistical methods should be described in detail. The MRI images should be presented to show the usefulness of MDF. The

followings are several concerns that the authors may wish to consider.

Reply: Thank you for your comment.

2) Specific comments

Major concerns:

1. In Radiomic analysis of Methods section, the authors described that the useful features were selected among 101 features in each sequence using algorithms, i.e., mutual information (MI), Fisher coefficient (Fisher) and classification error probability, which was combined with average correlation coefficients (POE + ACC and PA). These combinations led to the 30 highest discriminative power features in each sequence for further analysis. It is difficult, however, how the authors selected the 30 highest discriminants and calculated a probability from the 30 discriminants in combination. What is the POE + ACC and PA? Please explain the methods in detail and discuss what are expected to be the major determinants of MVI from the point of MDF.

Reply: Thank you for your comment. PA is the abbreviation of POE + ACC. MI, Fisher and PA are the screening methods of the MaZda software. Each of MI, Fisher, PA selects 10 texture parameters from the input data. Therefore, the maximum allowable number of input parameters is 30. And the 30 parameters can be further used for texture analysis by means of the B11 module, which can be executed in MaZda by Tools=>B11(Figure 1). We put a part of the screenshot from the MaZda manual above. And the screenshot from the MaZda manual about the details of MDF has been put below (Figure 2).

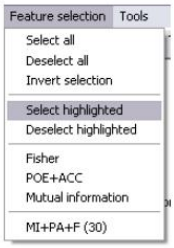

Fig. 1.5.7 The Feature selection menu items of the report window

The selected (checked on the list) texture parameters can be used as an input to feature selection/reduction procedures. There are two such procedures implemented in **MaZda**: one based on Fisher coefficient (a ratio of between-class to within-class variance) and the other that selects parameters that classify the given textures with the smallest arror and are least correlated with each other (Fig. 1.5.7). Each of them selects *10 texture parameters from the input list. These* procedures were previously the core modules of the **Convert** program.

The features of the tabs of the report windows can be saved to disk, one by one by using the **File=>Save report** menu item. The selected features only can be saved by means of **File=>Save selected** menu item.

The texture parameters selected, either manually or by means of the automated "Fisher"/"POE+ACC" procedures, can be further used for texture analysis by means of the **B11** program, which can be called from within **MaZda** by **Tools=>B11 analysis** menu item of the report window (Fig. 1.5.8). The maximum allowable number of input parameter to **B11** is 30.

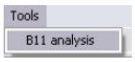

Figure 1

The goal of linear discriminant analysis is to find a linear transform matrix Φ such that the ratio of determinants

$$\frac{\left|\Phi^T \mathbf{C}_T \Phi\right|}{\left|\Phi^T \mathbf{C}_W \Phi\right|}$$

is maximized (Fukunaga 1991, Mao and Jain 1995). It can be proved that such a transform Ψ is composed of eigenvectors corresponding to largest eigenvalues of $\mathbf{C}_W^{-1}\mathbf{C}_T$. Transformation of original data by means of matrix Ψ

$$\mathbf{q}_i = \Psi^T(\mathbf{x}_i - \boldsymbol{\mu})$$

produces most discriminating features, MDF, (Swets and Weng 1996), such that $\mathbf{MDF}_i = [MDF_{i1}, MDF_{i2}, \ldots, mdf_{iNq}] = \mathbf{q}_i = [q_{i1}, q_{i2}, \ldots, q_{iNq}]$, i=1,2,...,M, where Nq is usually smaller than Nx.

To evaluate the usefulness of LDA to classes discrimination, the *linear separability* coefficient l_s (Mao and Jain, 1995) is

calculated, which is defined as the largest eigenvalue of $\mathbf{C}_T^{-1}\mathbf{C}_B$ (Gallinari et al 1991). As l_s changes from 0.0 to 1.0, the data set becomes more and more linearly separable. Similarly to PCA analysis, an LDA dimensionality factor is calculated, which is equal to the number of largest eigenvalues of $\mathbf{C}_T^{-1}\mathbf{C}_B$ whose sum is greater than 0.97.

Figure 2

2. Histogram features are included in the 101 features that were used to develop MDF consisting of 30 features. Why were histogram features separately subjected for the validation study? Please describe the reason to pick up histogram features. Furthermore, it would be helpful to understand the effect of MDF on the prediction of MVI that the authors present MRI images at AP and PVP for the cases with MVI+ and MVI-, which show similar histogram features but different MDF.

Reply: Thank you for your comment. The reason why we chose histogram parameters

was that it was a classic and basic image feature. Various research has demonstrated its good clinical use. For example, Histogram parameters have been used in quantitative analysis of MVI in clinical studies[1, 2]. Li et al. performed histogram analysis in intravoxel incoherent motion and the best parameter provided a sensitivity of 81% and a specificity of 85%[1]. Wang et al. used computational quantitative measures based on the maximum cross-sectional area to predict the MVI of a small HCC, but only in HBP images[2]. The AUC, sensitivity, and specificity of 0.91, 87 %, and 80 %, respectively. And MRI images of four cases with MVI+ and MVI- at AP and PVP were presented, which showed similar histogram features but different MDF (Page10, line268-270; Fig. 4 and Table S1).

1. Li H, Zhang J, Zheng Z, Guo Y, Chen M, Xie C, et al. Preoperative histogram analysis of intravoxel incoherent motion (IVIM) for predicting microvascular invasion in patients with single hepatocellular carcinoma. Eur J Radiol. 2018;105:65-71. doi: 10.1016/j.ejrad.2018.05.032,

2. Wang X, Zhang Z, Zhou X, Zhang Y, Zhou J, Tang S, et al. Computational quantitative measures of Gd-EOB-DTPA enhanced MRI hepatobiliary phase images can predict microvascular invasion of small HCC. European journal of radiology. 2020;133:109361.

3. The major purpose of this study is to show the higher accuracy of an algorithm using MDF in prediction of MVI comparing with a regular image diagnosis by radiologists. I believe that histogram features are considered to be what radiologists get from the information to reach their diagnosis. However, I am not sure that histogram features actually involve the information just enough for the diagnosis by radiologists. The efficacy of the algorithm should be directly compared with the diagnosis that was made by radiologists.

Reply: Thank you for your comment. The predictive power of MDFs derived from the radiomics analysis was better than that of all other histogram parameters (AUC: T1WI range from 0.52-0.68, T2WI range from 0.53-0.70, AP range from 0.54-0.69, PVP range from 0.50-0.74, EP range from 0.51-0.74 and HBP range from 0.52-0.65). Thus, we generated the ROC curves of MDF_{AP} and $MDF_{PVP,}$ which were independent predictors. The ROC curves of imaging features which statistically significant

differences were also generated. The results were compared using **the Delong test**. The MDF_{AP} and MDF_{PVP} had significantly higher AUCs than MTD, arterial rim enhancement and tumor margin ($P < 0.05$). However, there were no differences in AUCs between MDF_{AP}, MDF_{PVP} and peritumoral hypointensity in HBP ($P > 0.05$) (Page11-12, line300-311; Page11, line324-325; Fig. 6; Table S1)

Minor concerns:

1. In image analyses, the largest cross-sectional area was evaluated for MVI. Then, which section was evaluated in histology for MVI?

Reply: Thank you for your comment. Radiomic analysis was performed only on the largest cross-sectional area and two adjacent images of the tumor. There may be information loss compared to whole tumors which was evaluated in histology for MVI. But it is difficult to match the pathological sampling sites and MRI images not only derived from 2D ROIs but also 3D VOIs. Some articles used the method of the largest cross-sectional area for research. For example, Wang et al.[3] used computational quantitative measures to predict the MVI of a small HCC, which was based on the maximum cross-sectional area as well. The AUC, sensitivity, and specificity of 0.91, 87 %, and 80 %, respectively.

3. Wang X, Zhang Z, Zhou X, Zhang Y, Zhou J, Tang S, et al. Computational quantitative measures of Gd-EOB-DTPA enhanced MRI hepatobiliary phase images can predict microvascular invasion of small HCC. European journal of radiology. 2020; 133:109361.

2. Original T2WI should be presented without the coloration showing ROI.

Reply: Done.

Reviewer #2:

Scientific Quality: Grade B (Very good)

Language Quality: Grade A (Priority publishing)

Conclusion: Accept (General priority)

Specific Comments to Authors: I am applicated to review your valuable manuscript. I hope that some comments may be helpful to improve your manuscript. Your sample size was 113 including 73 MVI (+) cases and 40 MVI (-) cases. In such case, only seven variables allow to enter to multivariate logistic analysis to detect the risk factor to MVI (+) which may contribute to worse disease-free survival for cases. Readers may be understandable if you add logical explanation about your multivariate analysis process. Thank you and best regards.

Reply: Thank you for your comment. Firstly, the variables (including clinical factors, imaging features and MDFs in different sequences) were evaluated by the univariable logistic regression analysis. Secondly, multivariable logistic regression analysis with forward stepwise elimination includes the variables with a *P* value inferior to 0.05 at univariable analysis, i.e. the seven variables. Finally, three variables were entered into the final model as the risk predictors. And we add related content in the "Statistical Analysis" section. (Page9, line233-235)

Reviewer #3:

Scientific Quality: Grade C (Good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Minor revision

Specific Comments to Authors: Yueming Li et al. developed a radiomic analysis model based on preoperative magnetic resonance imaging (MRI) data to predict MVI in HCC. The overall research is relatively complete, but I have a few questions that I would like to discuss with the author.

1. In other articles, the identification methods of most independent predictors are univariate and multivariate COX regression analysis, and the author uses univariate and multivariate logistic regression analysis to identify. I hope the author can explain the advantages of using logistic regression analysis.

Reply: Logistic regression is a popular statistical method for studying the effects of covariates on binary outcomes. Univariate and multivariate logistic regression analysis were performed to build predictive models for MVI status (MVI + and MVI -) in our study. Cox regression builds a predictive model for time-to-event data.

2. The description of the method part is not specific, and the reader cannot repeat it completely.

Reply: It could be repeated following the steps and combined with the manual of MaZda [4,5]. And we have added some modifications to the method. In addition, this software has often been used in various studies in recent years [6,7].

4. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda--a software package for image texture analysis. Computer methods and programs in biomedicine. Comput Methods Programs Biomed 2009;94(1):66-76.

5. Strzelecki M, Szczypinski P, Materka A, Klepaczko AJNI, A MiPR. A software tool for automatic classification and segmentation of 2D/3D medical images. Nucl Instrum Methods Phys Res A 2013;702(2):137-40.

6. Ye R, Weng S, Li Y, et al. Texture Analysis of Three-Dimensional MRI Images May Differentiate Borderline and Malignant Epithelial Ovarian Tumors. Korean J Radiol. 2021;22(1):106-117. doi:10.3348/kjr.2020.0121.

7. Wang R, Su Y, Mao C, Li S, You M, Xiang S. Laser lithotripsy for proximal ureteral calculi in adults: can 3D

CT texture analysis help predict treatment success? [published online ahead of print, 2020 Nov 19]. Eur Radiol. 2020;10.1007/s00330-020-07498-x. doi:10.1007/s00330-020-07498-x.

3. The author uses R software, but the specific method and R package are not cited.

Reply: The five-fold cross-validation was performed using the "caret" package, and nomogram was used as a graphical representation using the "rms" package (R software version 4.0.2, http://www.r-project.org). (Page 9, line236-238)

4. The results of independent prognostic factors should be visualized with pictures.

Reply: Thanks for your suggestion. The independent predictive factors were integrated into a **nomogram** by the multivariate logistic regression analysis in figure 6B.

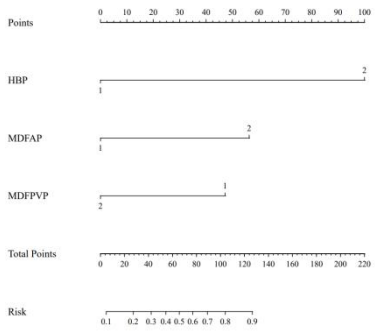

figure 6B

5. The results of the 5-fold cross-validation should also be displayed with pictures.

Reply: Thanks for your suggestion. The ROC curve of the 5-fold cross-validation was added in figure 6A.

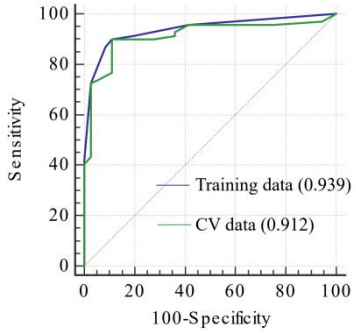

figure 6A

6. The author only constructed a model based on 113 patients. The model was not verified by external independent data. I am very worried about its accuracy and whether it can be applied to all patients.

Reply: Although the small sample size is a limitation of our study, it took three years for us to select 113 cases according to the inclusion standards and exclusion standards. To reduce the impact of the small sample size, internal validation of 5-fold cross-validation was also applied via R software (version4.0.2). External validation by larger datasets from other centers is still needed. We will continue to select more cases and do further study in the future.

7. Part of the picture results can be merged, instead of putting a small picture in a whole figure.

Reply: We have added and merged some pictures in Figure 4); Figure 6).

8. The language part of the article still needs some polishing.

Reply: This manuscript was polished by English professionals with a service letter. According to your comment, our revised manuscript has been sent to be embellished by English professionals again.

###**Science editor's comments:**

Language Quality: Grade B (Minor language polishing)

Scientific Quality: Grade C (Good)

1. Please revise the Methods section to explain the radiomics analysis section in detail; cite specific methods and R packages, etc.

Reply: More details could be from the manual of MaZda [4,5]. And we have added some modifications to the method. In addition, this software has often been used in various studies in recent years [6,7]. The five-fold cross-validation was performed using the "caret" package, and nomogram was used as a graphical representation using the "rms" package (R software version 4.0.2, http://www.r-project.org). (Page 9, line236-238)

4. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda--a software package for image texture analysis. Computer methods and programs in biomedicine. Comput Methods Programs Biomed 2009;94(1):66-76.

5. Strzelecki M, Szczypinski P, Materka A, Klepaczko AJNI, A MiPR. A software tool for automatic classification and segmentation of 2D/3D medical images. Nucl Instrum Methods Phys Res A 2013;702(2):137-40.

6. Ye R, Weng S, Li Y, et al. Texture Analysis of Three-Dimensional MRI Images May Differentiate Borderline and Malignant Epithelial Ovarian Tumors. Korean J Radiol. 2021;22(1):106-117. doi:10.3348/kjr.2020.0121

7. Wang R, Su Y, Mao C, Li S, You M, Xiang S. Laser lithotripsy for proximal ureteral calculi in adults: can 3D CT texture analysis help predict treatment success? [published online ahead of print, 2020 Nov 19]. Eur Radiol. 2020;10.1007/s00330-020-07498-x. doi:10.1007/s00330-020-07498-x

2. Results of independent prognostic factors should be visualized using pictures.

Reply: Thanks for your suggestion. The independent predictive factors were integrated into a **nomogram** by the multivariate logistic regression analysis in figure 6B.

3. Furthermore, whether or not the histogram features actually contain just enough information to be diagnosed by a radiologist, the efficacy of the algorithm should be directly compared with the radiologist's diagnosis.

Reply: Thank you for your suggestion. The ROC curves of imaging features which

were statistically significantly different were also generated. The results were compared using **the Delong test**. The MDF_{AP} and MDF_{PVP} had significantly higher AUCs than MTD, arterial rim enhancement, and tumor margin ($P < 0.05$). However, there were no differences in AUCs between $MDF_{AP,}$ MDF_{PVP} and peritumoral hypointensity in HBP ($P > 0.05$) (Page9, line233-235; Page11-12, line300-307; Page12, line320-321; Figure 6; Table S1/S2)

###Company editor-in-chief's comments:

1. Before final acceptance, uniform presentation should be used for figures showing the same or similar contents; **for example, "Figure 1Pathological changes of atrophic gastritis after treatment. A: ...; B: ...; C: ...; D: ...; E: ...; F: ...; G: ...**". Please provide decomposable Figures (in which all components are movable and editable), organize them into **a single PowerPoint file**. Please authors are required to provide standard three-line tables, that is, only the top line, bottom line, and column line are displayed, while other table lines are hidden. The contents of each cell in the table should conform to the editing specifications, and the lines of each row or column of the table should be aligned. Do not use carriage returns or spaces to replace lines or vertical lines and do not segment cell content. In order to respect and protect the author's intellectual property rights and prevent others from misappropriating figures without the author's authorization or abusing figures without indicating the source, we will indicate the author's copyright for figures originally generated by the author, and if the author has used a figure published elsewhere or that is copyrighted, the author needs to be authorized by the previous publisher or the copyright holder and/or indicate the reference source and copyrights. Please check and confirm whether the figures are original (i.e. generated de novo by the author(s) for this paper). If the picture is 'original', the author needs to add the following copyright information to the bottom right-hand side of the picture in PowerPoint (PPT): Copyright ©The Author(s) 2022

Reply: Done. The single PowerPoint file of the original images was provided.

ROUND 2

Dear editor and reviewer:

Thank you very much for your comments.

We apologized that we did not respond appropriately to the concerns you raised in the first-round review. In general, this study is aimed to explore a radiomic analysis model for the prediction of MVI in HCC. The AUC of **final radiomic nomogram model include MDF_{AP}, MDF_{PVP} and peritumoral hypointensity on HBP** was 0.939. It outstrips the single feature of peritumoral hypointensity on HBP by visual inspection of radiologists **(AUC: 0.849)**. Even though there is no significant difference in predictability of MVI between MDF_{AP}/MDF_{PVP} **(AUC: 0.845/0.881)** and peritumor hypointensity in the images of hepatobiliary phase alone, we think the MDF_{AP}/MDF_{PVP} still provide additional diagnostic value in the fusion radiomic nomogram **(AUC: 0.939)**. Once again, we carefully understand your concerns and review the relevant literature. We have revised our manuscript according to your comments. Here below is our reply to your and the reviewer's comments. We have also modified some corresponding expressions in our manuscript. Hope to help readers better understand the purpose and method of this study.

Major concerns:

1. For the concern in my first review that it is hard to understand what kind of image features were extracted to predict microvascular invasion (MVI) in this study, the authors replied by simply attaching the instruction of the software that the authors used. Readers may not be a user of MaZda software. The authors must transfer the concept of image analyses to readers not by the instruction but by regular expressions. Is MDF expected to predict MVI at the image sections that were not directly evaluated in image analyses? If MDF represents some concept of image feature, the authors should dictate it and explain the concepts of MDF_{AP} and MDF_{PVP}. If it is impossible to follow the calculation steps and dictate the concept

of each MDF as an output of artificial intelligence, explicitly describe it.

Reply: Thank you for comments. We apologize that there is not clear for us to explain the feature of MDF in our first-round review respond. Microvascular invasion (MVI) has been widely recognized as one of the potential predictors of HCC recurrence. It can only be diagnosed pathologically at present instead of naked eye by radiologists. So we are pleased to find in our study that noninvasive MRI radiomic model of MDF values and imaging biomarkers may be useful to make a preoperative prediction of MVI in patients with primary HCC. The most discriminant factor (MDF) is a comprehensive variable for discrimination and represent **a linear discriminant analysis** of these input 30 features which were selected by mutual information (MI), Fisher coefficient (Fisher) and probability of classification error and average correlation coefficients (POE+ACC). So, the linear transformation of original data by means of special matrix produces most discriminating features (MDF). We can use the statistical B11 module (a plug-in of Mazda software) to handle complex matrix transformations. We have also modified some corresponding expressions in **page 8 paragraph 2**, to help readers understand what we mean. Please contact us if you have any questions.

2. In my first review, I believed that histogram features are the digital expression of diagnosis made by a radiologist. However, now I understand that those two are different. In Figure 4, the authors presented that histogram features were similar between cases A and B, and between cases C and D. The impressions for MVI between cases A and B and between cases C and D are clearly different by visual inspection. I believe that no radiologist would make diagnose of MVI-negative in cases B and D. I asked in my first review that the efficacy of the algorithm should be directly compared with the diagnosis that was made by radiologists. Furthermore, if there is no significant difference in predictability of MVI between MDF_{AP}/MDF_{PVP} and peritumor hypointensity in the images of hepatobiliary phase, what is the efficacy to calculate the MDFs in clinic. If Figure 4 are

representative cases showing the efficacy of MDFs, I am reluctant to say that calculation of MDF is waste of time for MVI prediction.

Reply: Your understanding is correct. Histogram is a classical first-order radiomic features, which has been widely used in various fields of medical research. Therefore, we compared MDF with histogram in order to explore the efficacy of both in predicting MVI. Although the qualitative image features for MVI between cases A and B and between cases C and D are different by visual inspection, it is still worth studying whether there are differences in the quantitative MDF values before statistical analysis in our study. There is no significant difference in predictability of MVI between MDF_{AP}/MDF_{PVP} **(AUC: 0.845/0.881)** and peritumor hypointensity in the images of hepatobiliary phase **(AUC: 0.849)** alone, but we think the MDF_{AP}/MDF_{PVP} still provide additional diagnostic value in the fusion radiomic nomogram which includes MDF_{AP}, MDF_{PVP} and peritumoral hypointensity on HBP **(AUC: 0.939)**. The efficacy of peritumor hypointensity in hepatobiliary phase in predicting MVI has been confirmed by many studies, but its specificity is high, while its sensitivity is still insufficient. So, the combined MDF values may improve the prediction efficiency of MVI.

Also, it is still worth studying whether there are differences in the quantitative MDF values to predict HCC MVI before statistical analysis in our study although the diameter of tumors in Figure 4 in cases B and D is larger and more typical than that in cases A and C. The reason that we choose the images is just to show the features easily for the readers. If necessary, we can also amend the images in cases B and D according to your comments.

Minor concerns:

1. Comma, Space, etc. are not correct. Many two words are stuck in one. Still the abbreviations of POE and ACC are not spelled out. If PA = POE + ACC, then the expression of “POE + ACC and PA” may not be appropriates

Reply: Thank you for your kindly comment. All the errors about comma, space, etc.

have been carefully checked. Also, it's an mistake of "PA", and.we have deleted "and PA" and modified the corresponding expressions in our manuscript.

Other comment:

1. Please provide the fund document Supported by This study has supported by Joint Funds for the Innovation of Science and Technology, Fujian Province (CN) (Award Number: 2019Y9125)

Reply: Thank you for your kindly comment. The fund document of this study has also been resubmitted.

Sincerely yours,

Dr. Yueming Li

fjmulym@163.com

April 5, 2022