MAYO CLINIC

**Lewis R. Roberts, M.B., Ch.B., Ph.D.**
Division of Gastroenterology and Hepatology
200 First Street SW
Rochester, Minnesota 55905
Tel: 507-266-4720
Fax: 507-284-0762
Email: Roberts.Lewis@mayo.edu

December 18, 2017

*World Journal of Gastroenterology*

Dear Editors:

We thank the *World Journal of Gastroenterology* for reviewing our manuscript entitled, "Pre-transplant BALAD and BALAD-2 in predicting hepatocellular carcinoma patients recurrence and survival" by Wongjarupong et al. We have reviewed the comments and have considered them carefully. We have responded to each comment or request in detail. We trust our responses will prove acceptable. The point-by-point response to reviewers' and editors' comments can be found below.

**Reviewer 1**

**Comment 1:** (Material and methods) What is the definition of the elevation of each tumor marker in BALAD score? It should be clarified in the figure 1. Is it the same with that in figure 2? If so, how was the cutoff level calculated?

**Response:** Thank you for this comment. We have added the definition of the elevation of each tumor biomarker in BALAD score in the **Table 1** (changed from **Figure 1**) as follow:

"Defined by AFP >400 ng/mL, AFP-L3 >15%, and DCP >100 ng/mL"

**Comment 2:** (Material and methods) What is the definition of HCC recurrence after liver transplantation?

**Response:** We have added the definition of the HCC recurrence in the Methods section on Page 7 as follows:

"HCC recurrence was defined by the presence of new malignant masses seen on imaging, either intrahepatic or extrahepatic metastases, as assessed by the radiologist."

**Comment 3:** The formulas for GALAD and GALAD-z should be clarified.

**Response:** The formulas for the GALAD score and GALAD-z calculation have been added as Table 1c.

**Comment 4:** (Discussion section, the 6th paragraph) The authors described that 'with and without HCC recurrence, We compared …' should be 'with and without HCC recurrence. We compared …'

**Response:** We thank the reviewer for this comment. We have changed the sentence as suggested in the Discussion section page 13.

**Comment 5:** Figure 1 is not a figure but a table.

**Response:** We have changed **Figure 1** to **Table 1a** and **1b**.

**Reviewer 2**

**Comment 1:** The authors aimed at validating BALAD and BALAD 2 scores but instead they modified the original scores by re-weighting their components in a different formula. There is no true validation. If the original aim is to be pursued, BALAD and BALAD 2 scores should be individually calculated for each patient and be included as such in the multivariate Cox's regression analysis, where other well-known predictors of HCC recurrence may be controlled (diameter of the main nodule, number of nodules, microvascular invasion, macrovascular invasion, histological tumor differentiation…). In contrast, if the authors aim at re-weighting the score formula, it would be mandatory to validate the new resulting scores in an external cohort (otherwise external validity is compromised). None of these options are successfully accomplished in the present version of the manuscript.

**Response:** We thank the reviewer for this essential point. We originally aimed to validate the BALAD score and BALAD-2 class. We had left out the multivariate analysis due to space constraints, but have now added the analysis with diameter of the largest tumor and neutrophil-lymphocyte ratio back as suggested by the reviewer in Table 3b and 3c. The following statement of the multivariate analysis result was also added to the Results section on page 11.

"A multivariate model of diameter of the largest tumor and neutrophil-lymphocyte ratio of more than 4 with BALAD and BALAD-2 was created (**Table 3b and 3c**). The risk of recurrence was 1.53 (1.17-2.01) per increase of 1 in the BALAD score and 1.42 (1.05-2.03) per increase of one BALAD-2 class. The risk of death was 1.57 (1.27-1.96) per increase of 1in the BALAD score and 1.37 (1.07-1.76) per increase of 1 BALAD-2 class."

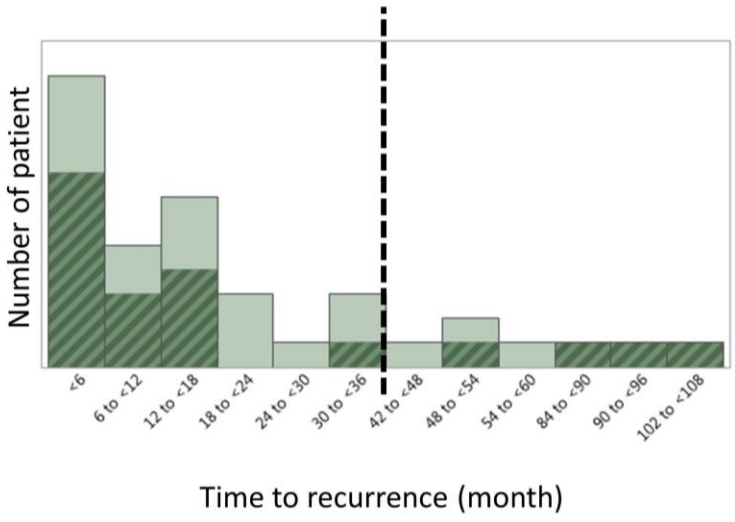
We would like to introduce the S-LAD score as the BALAD does not perform as well in the liver transplant setting as the S-LAD. However, we agree that further validation of the S-LAD in additional cohorts will be the subject of future studies. We have added the need for further validation to the Conclusion.

**Comment 2:** The cohort of patients is not consecutive (only 113 patients out of 299 were included because of missing data). Clearly the study is based in a subpopulation with more advanced HCC (with unusually high tumor recurrence rates), which does not represent the entire HCC transplanted population from the Mayo Clinic. Internal validity may be therefore compromised. This fact becomes apparent in the discussion where it can be read: "Thirty-eight of 113 patients (33.6%) with available serum had recurrence. However, when considering all HCC patients who underwent liver transplant during the same period, 43 of 299 patients (14.4%) had recurrence".

**Response:** We acknowledge this point made by the reviewer as a major limitation of the study and have emphasized this in the discussion. As indicated in the discussion, samples from patients without recurrence appeared to have been more frequently requested from the biobank than samples from patients with recurrence, which at least partially led to the apparently high recurrence rate seen in our cohort. To control for this potential bias, we compared the baseline characteristics of patients who had no recurrence and available biomarker results to those of patients who had no recurrence but no biomarker results (**Supplemental data 2**). Encouragingly, patients in both groups had similar baseline characteristics, suggesting that there was no systematic bias in the patients who had biomarker results compared to those without biomarker results.

**Comment 3:** Some authors have suggested that HCC after 3 years post-transplant may not be considered recurrence, but “de novo” HCC. In the present study there was a prolonged surveillance after transplantation and a significant proportion of patients were transplanted with chronic hepatitis C. Since the study was performed in the pre-DAA era, it may well be that some of the patients with late HCC recurrence were actually new HCCs within a recurrent hepatitis C. I would recommend removing late HCC recurrence from the analysis and controlling hepatitis C status in the multivariate analysis.

**Response:** We thank the reviewer for this insightful question. There were 31 patients who had recurrence within 3 years, with 2 patients developing recurrence within 1 month of liver transplantation, and 7 patients who developed recurrence more than 3 years after liver transplantation. The latter 7 patients developed recurrence at 42, 49, 50, 59, 89, 94, and 107 months, respectively. Below is a bar graph showing recurrence events classified by time to recurrence in 6 month periods. The dark green color bars represent patients with HCV infection.




| Level | Count | HCV |
|---|---|---|
| <6 | 12 | 8 |
| 6 to <12 | 5 | 3 |
| 12 to <18 | 7 | 4 |
| 18 to <24 | 3 | 0 |
| 24 to <30 | 1 | 0 |
| 30 to <36 | 3 | 1 |
| 42 to <48 | 1 | 0 |
| 48 to <54 | 2 | 1 |
| 54 to <60 | 1 | 0 |
| 84 to <90 | 1 | 1 |
| 90 to <96 | 1 | 1 |
| 102 to <108 | 1 | 1 |
| Total | 38 | 20 |

Of the 7 patients who developed recurrence more than 3 years after transplant, 4 patients had HCV infection (57%) whereas in the whole cohort of 113 patients, 66 had HCV (58%). We explored the HCV treatment status of the four patients with HCV who developed late HCC recurrence. Two patients had recurrent HCV within one year after the transplant, one patient had undetectable HCV RNA after the transplant throughout the follow-up period, and one patient had unknown HCV infection status. The univariate Cox proportional hazard ratio for recurrence of HCC (both early and late recurrence) in patients with HCV was 0.78 (0.41-1.49, $p$=0.46) compared to patients without HCV.

We also calculated the risk of recurrence after excluding the late recurrences that occurred more than 3 years after transplant. The results are shown in the Table below. This table is added as **Supplemental data 2**.

| Variable | Hazard ratio for recurrence | | Hazard ratio for early recurrence | |
|---|---|---|---|---|
| | HR (95%CI) | *P* value | HR (95%CI) | *P* value |
| MELD score (per point) | 1.03 (0.98-1.09) | 0.26 | 1.02 (0.96-1.08) | 0.43 |
| Maximal tumor size at time of transplant (per cm) | 1.27 (1.04-1.56) | 0.02 | 1.26 (1.009-1.57) | 0.039 |
| Diameter of the largest tumor at time of transplant | 1.001 (0.73-1.37) | 1.00 | 1.08 (0.76-1.43) | 0.64 |
| Neutrophil lymphocyte ratio >4 | 2.24 (1.17-4.26) | 0.02 | 2.65 (1.30-5.63) | 0.008 |
| Hypothyroidism | 1.26 (0.55-2.85) | 0.59 | 1.40 (0.52-3.19) | 0.48 |
| BALAD components | | | | |
| - Albumin (per g/dL) | 0.75 (0.41-1.38) | 0.36 | 0.73 (0.37-1.41) | 0.36 |
| - Bilirubin (per mg/dL) | 1.03 (0.98-1.09) | 0.21 | 1.02 (0.95-1.08) | 0.47 |
| - AFP: >400ng/mL | 2.42 (1.18-5.00) | 0.02 | 3.10 (1.34-6.55) | 0.005 |
| - AFP-L3 >15% | 1.86 (0.98-3.52) | 0.056 | 1.92 (0.95-3.94) | 0.07 |
| - DCP > 1.2 ng/mL | 2.83 (1.42-5.61) | 0.003 | 3.73 (1.74-8.93) | 0.0006 |
| BALAD Score | | | | |
| 0 | Reference | | Reference | |
| 1 | 0.70 (0.20-2.47) | 0.58 | 1.41 (0.28-7.01) | 0.66 |
| 2 | 1.18 (0.37-3.75) | 0.78 | 1.75 (0.36-8.44) | 0.46 |
| 3 | 1.99 (0.62-6.36) | 0.24 | 3.78 (0.82-17.53) | 0.055 |
| 4 | 2.97 (0.84-10.58) | 0.09 | 5.22 (1.01-27.01) | 0.04 |
| 5 | 5.02 (0.92-27.54) | 0.06 | 15.63 (2.10-116.17) | 0.02 |
| BALAD Score (per increase of 1) | 1.48 (1.15-1.91) | 0.002 | 1.66 (1.24-2.22) | 0.0006 |
| BALAD-2 Score | | | | |
| 1 | Reference | | Reference | |
| 2 | 0.41 (0.12-1.32) | 0.13 | 0.57 (0.14-1.88) | 0.35 |
| 3 | 1.53 (0.66-3.54) | 0.32 | 1.63 (0.64-4.42) | 0.31 |
| 4 | 2.17 (0.90-5.25) | 0.09 | 2.41 (0.89-6.75) | 0.08 |
| BALAD-2 Score (per increase of 1) | 1.45 (1.06-1.98) | 0.02 | 1.46 (1.04-2.07) | 0.03 |
| Within Milan criteria at diagnosis | 1.69 (0.84-3.41) | 0.14 | 2.16 (0.97-4.48) | 0.058 |
| Within UCSF criteria at diagnosis | 1.85 (0.85-4.05) | 0.12 | 2.36 (0.94-5.23) | 0.07 |
| Within Milan criteria at transplant | 1.24 (0.59-2.62) | 0.57 | 1.65 (0.72-3.47) | 0.22 |
| Within UCSF criteria at transplant | 0.33 (0.05-2.43) | 0.28 | 0.88 (0.14-2.91) | 0.86 |
| z-GALAD | 1.12 (1.03-1.21) | 0.006 | 1.14 (1.05-1.23) | 0.003 |
| GALAD score | 3.01 (1.14-7.91) | 0.03 | 3.74 (1.41-10.61) | 0.008 |
| AFP model cutoff > 2 (explant) | 2.82 (1.47, 5.41) | 0.002 | 3.51 (1.78-6.73) | 0.0005 |
| AFP model (per increase of 1, explant) | 1.42 (1.20, 1.68) | <0.001 | 1.48 (1.24-1.75) | <0.001 |

In addition, we addressed this difference of early and overall recurrence in the Results section on page 11 as follows:

“In addition, the hazard ratios for early recurrence were also calculated. Early recurrence was defined as recurrence occurring within 36 months after transplant. Of the 38 patients with any recurrence, 31 had early recurrence. The BALAD score had better performance for early than overall recurrence with a HR of 1.66 (1.24-2.22) per each unit increase of BALAD score, whereas the BALAD-2 class had similar performance for both recurrence outcomes with a HR of 1.46 (1.04-2.07) per increase of 1 class (**Supplemental data 2**).”

**Comment 4:** The methodology for sample size calculation is not described.

**Response:** We apologize for the lack of clarity regarding the study design. The study was designed with assistance from a PhD level and Master’s level statisticians. The total sample size was limited by the number of samples available for analysis. For the multivariate analyses, we limited the number of variables analyzed based on the number of recurrence or death events in the cohort, using the recommended ration of 10-20 events per variable to reduce the likelihood of false positive results.

**Comment 5:** “Tumor size and tumor number were also determined from the most recent imaging studies prior to the transplant.” Why not from the explanted liver?

**Response:** Thank to the reviewer for raising this interesting point. Our intent in performing this study was to determine whether a score based on the pre-transplant AFP, AFP-L3 and/or DCP would have utility in determining whether individuals were eligible for liver transplant or not as per the currently used AFP model. This led to our preference of using information that was available prior to the transplant.

To address the reviewer’s question, we also calculated the Cox proportional hazard ratios for the explant diameter of the largest tumor. The hazard ratios were 1.38 (1.14-1.65, $p$=0.0008) per 1 cm increase in tumor size for recurrence and 1.25 (1.08-1.45, $p$=0.003) per 1 cm increase in tumor size for death. The Cox proportional hazard ratios for the pre-transplant diameter of the largest tumor were 1.27 (1.04-1.56, $p$=0.02) and 1.21 (1.03-1.41, $p$=0.02), respectively, which were comparable to the results from the explant data.

**Comment 6:** In results it can be read: “The transplant selection criteria for the HCC patients during the study period were mainly based on the Milan criteria and extended criteria, the UCSF criteria, in some patients”. This information should be transferred to the methods’ section. In addition, it is unclear in which situations were UCSF criteria allowed. Please clarify.

**Response:** We thank the reviewer for this comment. We have moved the sentences to the first paragraph of the Methods section on page 7. Regarding the application of the UCSF criteria instead of the Milan criteria, we have further explored our data for this interesting issue. There were 17 patients that were within the UCSF criteria but not the Milan criteria at the time of transplant. Thirteen of the seventeen patients underwent liver transplant after 2002, the year that the UCSF criteria were published by Yao et al.. We suspect that the expansion of the selection criteria was due to the increasing evidence at the time that the Milan criteria could be safely expanded while substantially preserving the survival benefits of liver transplantation in patients with HCC.

**Comment 7:** In addition, there were some patients beyond UCSF criteria both at listing and at the time of transplant (15% and 7%). I understand that this information comes from pre-

transplant imaging assessment. In most transplant institutions, these patients would have been removed from the waiting list in line with current guidelines. Please clarify the reason for not excluding these patients.

**Response:** In the UNOS region that included Florida at the time the blood samples were obtained, the organ availability for liver transplantation was such that if the program was willing to use less than ideal organs, these were available for some patients who were beyond UCSF criteria. Thus, of the 8 patients who were outside the UCSF criteria at the time of transplant, 7 patients underwent transplant at Mayo Clinic Florida, while only one 1 patient underwent transplant at Mayo Clinic Rochester.

**Comment 8:** Was there any protocolized follow-up after LT to detect tumor recurrence? Please describe.

**Response:** Patients underwent HCC surveillance per the same protocol for all three Mayo Clinic sites, with. CT of the abdomen and chest along with serum AFP test at 4, 8, 12, 18, and 24 months post-transplant. The Results section has been modified on page 7 as follows:

“For surveillance for post-transplant HCC recurrence, patients underwent CT scan of the abdomen and chest along with serum AFP at 4, 8, 12, 18, and 24 months post-transplant.”

**Comment 9:** Describe the policy for prioritization within the waiting list and the length within the waiting list of the included patients.

**Response:** We have added the waiting list time at the end of the first paragraph in the Results section on page 9 as follows:

“The median waiting time for the included patients was 2.8 (range 0-20) months.”

In addition, the Standard UNOS policies operative at the time of the study in the respective UNOS regions serving Florida and Minnesota were followed. Our study included patients from 2000 to 2008. The MELD score was applied for prioritization beginning in 2002. The UCSF criteria was proposed by Yao et al. in 2002 and was widely applied afterward. Apart from the above, there was no major change of organ allocation for HCC patients until 2015, when the exception scores for HCC were modified.

**Comment 10:** Regarding locoregional bridging therapies within the waiting list, please provide indication criteria.

**Response:** Most patients with intermediate stage disease beyond the Milan criteria received locoregional treatment with transarterial chemoembolization. We have also added this to the Methods section on page 7 as follow:

“Most patients with intermediate stage disease beyond Milan criteria received locoregional treatment with transarterial chemoembolization.”

**Comment 11:** The term “maximal tumor size” is confusing. Do the authors mean the diameter of the largest nodule or total tumor size (sum of the diameters of each detected nodule)?

**Response:** Thank you for the reviewer comment. We have changed the term “maximal tumor size” to “diameter of the largest tumor” as suggested.

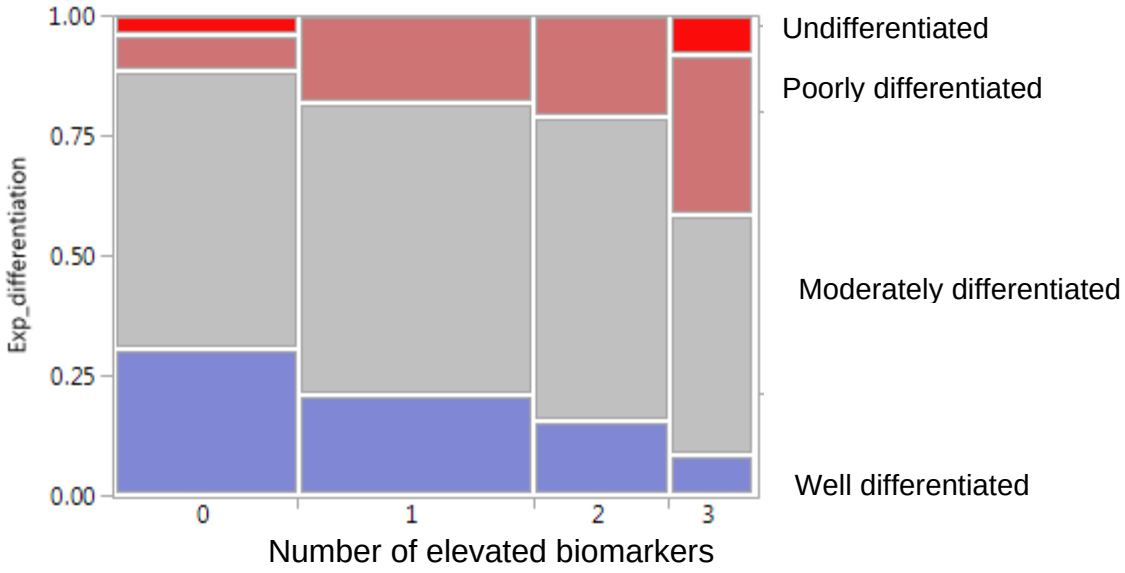
**Comment 12:** Information about microvascular invasion and histological tumor differentiation is missing. These histological features are tightly associated with tumor recurrence. The authors should consider this information for the analysis. It would be interesting to know the interaction between BALAD score and histological features of HCC (if any).

**Response:** We would like to thank the reviewer for this point. We have limited the number of variables analyzed by multivariate analysis to avoid false positives due to the relatively small number of patients with recurrence or death. In addition, we elected not to include information that we did not have at the time of transplant, i.e. explant data, in the prediction model.

Regarding the interesting question about tumor differentiation from the reviewer, we have analyzed the correlation between the number of elevated tumor biomarkers and tumor differentiation, as well as between the BALAD score and tumor differentiation in mosaic plots and in the tables below. Interestingly, we found a non-significant trend towards correlation between tumor differentiation and number of elevated tumor biomarkers (using the cut-off values for the BALAD score) and the BALAD score [IA and IB]. We have added this finding to the Results section on pages 9-10 and Supplemental data 1a and 1b as follows:

"According to the explant pathology reports, there were 19, 53, 16, and 2 patients with well-, moderately-, poorly-, and undifferentiated tumors, respectively. There were 23 patients with no report of tumor differentiation. The correlations of the number of elevated tumor biomarkers according to the BALAD score cut-off with the BALAD scores are shown in **Supplemental data 1**. There was no correlation between number of elevated tumor biomarkers ($p$=0.34), or BALAD score ($p$=0.28) with tumor differentiation."

**IA.** Mosaic plot of tumor differentiation by number of elevated biomarker (AFP, AFP-L3, DCP)

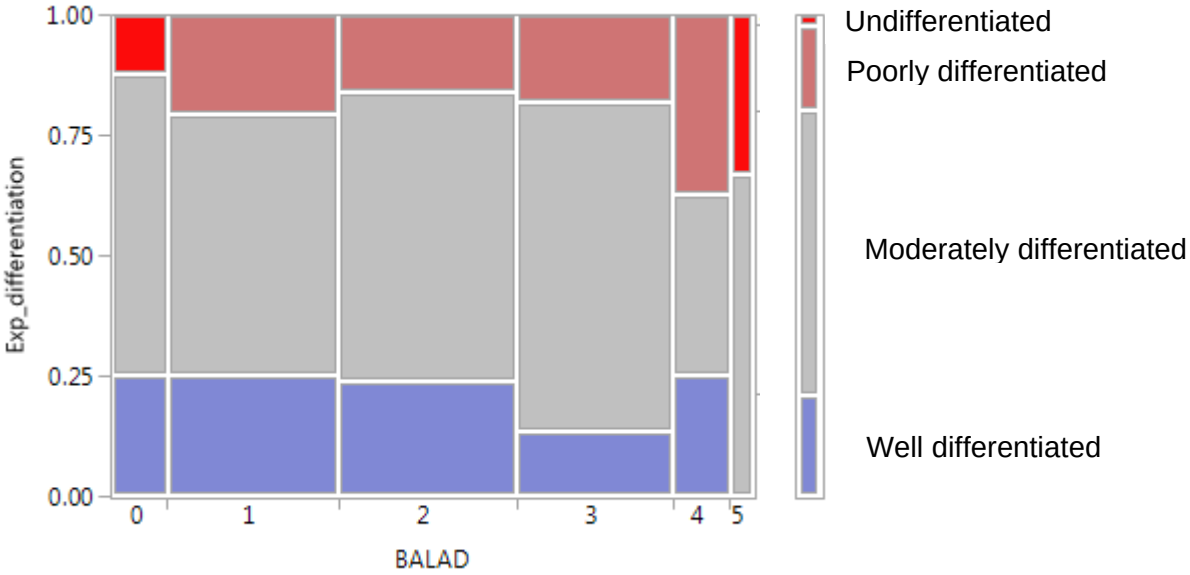



**IB.** Table of counts of tumor differentiation by number of elevated biomarker (AFP, AFP-L3, DCP)

| Number of elevated biomarker | Well differentiated | Moderately differentiated | Poorly differentiated | Undifferentiated | Unknown | Total |
|---|---|---|---|---|---|---|
| 0 | 8 | 15 | 2 | 1 | 14 | 40 |
| 1 | 7 | 20 | 6 | 0 | 7 | 40 |
| 2 | 3 | 12 | 4 | 0 | 1 | 20 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3 | 1 | 6 | 4 | 1 | 1 | 13 |
| Total | 19 | 53 | 16 | 2 | 23 | 113 |

**IIA.** Mosaic plot of tumor differentiation by BALAD score




**IIB.** Table of counts of tumor differentiation by BALAD score

| Number of elevated biomarker | Well diff 1 | Mod diff 2 | Poor diff 3 | Grade 4 | Unknown | Total |
|---|---|---|---|---|---|---|
| 0 | 2 | 5 | 0 | 1 | 6 | 14 |
| 1 | 6 | 13 | 5 | 0 | 7 | 31 |
| 2 | 6 | 15 | 4 | 0 | 8 | 33 |
| 3 | 3 | 15 | 4 | 0 | 1 | 23 |
| 4 | 2 | 3 | 3 | 0 | 1 | 9 |
| 5 | 0 | 2 | 0 | 1 | 0 | 3 |
| Total | 19 | 53 | 16 | 2 | 23 | 113 |

**Comment 13:** The conclusion “The BALAD score and BALAD-2 class is valid to predict recurrence and death in HCC patients with liver transplant” is not supported by the findings of the study. As outlined above the model was not analyzed as such, but instead was re-weighted and not externally validated.

**Response:** We agreed with the reviewer regarding the conclusion statement. Patients who underwent liver transplant regained their liver function which reduced the predictive value of the bilirubin and albumin for both recurrence and death outcomes. We have introduced the new model for transplant patients and emphasized the predictive role of the tumor biomarkers which represent tumor biology. We have changed the Discussion on page 15 as follows:

“In conclusion, the combination of the three biomarkers used in the BALAD score along with maximal tumor diameter (S-LAD) was the most predictive model for recurrence and death outcomes for HCC patients receiving liver transplants. However, validation of this new S-LAD model is warranted. Unlike the performance for other HCC treatment modalities, the BALAD

score and BALAD-2 class are less predictive for recurrence and death in HCC patients with liver transplant, presumably because liver function is restored after liver transplantation. ”

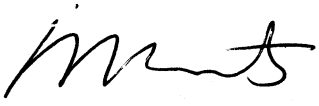
**Reviewer 3**

**Comment 1:** The BALAD models were designed to predict survival in patients with HCC. There are two elements to BALAD reflecting the major prognostic features for HCC - liver function (as measured by Bilirubin and Albumin) and the tumor-related factors (LAD i.e. the biomarkers). In patients undergoing liver transplantation, there is a ‘new’ liver and there is, therefore, no reason to believe that the preoperative liver function will impact on overall survival/outcome – other than, perhaps, as some indication of ‘overall health/performance status’. This has been well documented in studies that used just ALB and BILI (‘ALBI’) as a measure of liver function. In the initial publication (Johnson et al., J Clin Oncol. 20;33(6):550-8 2015) it was shown that ALBI had a major influence on survival in all therapeutic situations, except liver transplantation where there was no impact. Thus BALAD scores would appear, ‘a priori’, to be inappropriate in the transplant setting. I would suggest therefore that the paper just focuses on the biomarkers and their combination with tumour size (S-LAD). If the authors wish to retain the section on BALAD then this should not be the main feature of the paper but included to demonstrate the validity of the above argument that is to suggest that its application in the transplant setting is inappropriate.

**Response:** We agree with the reviewer. We have now refocused our study to introduce the new S-LAD model and also changed the name of our study to “Pre-transplant tumor biomarkers are more important in predicting hepatocellular carcinoma recurrence and survival compared to the BALAD models”. We also have added the suggested reference from Johnson et al. for this supportive statement in the Discussion section on Page 13. In addition, we would like to introduce the S-LAD score as the BALAD does not perform as well in the liver transplant setting as the S-LAD. We agree that further validation of the S-LAD in additional cohorts is needed and this will be the subject of future studies. We have modified the Conclusion section on page 14 accordingly.

We thank the editors and reviewers for their insightful comments. We appreciate all the reviewers’ suggestions which have improved our study. We believe the manuscript is stronger for them and we hope that our responses and manuscript modifications will prove satisfactory upon review.

Sincerely,



**Lewis R. Roberts, M.B., Ch.B., Ph.D.**
Professor of Medicine and Consultant
Director, Hepatobiliary Neoplasia Clinic
Division of Gastroenterology and Hepatology
Mayo Clinic, Rochester, MN