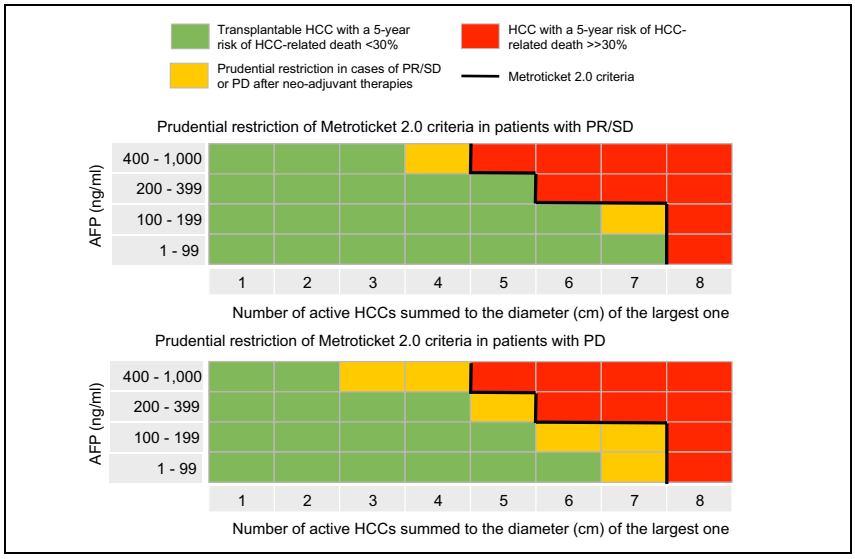
Research Article
Liver Transplantation

JOURNAL
OF HEPATOLOGY

Including mRECIST in the Metroticket 2.0 criteria improves prediction of hepatocellular carcinoma-related death after liver transplant

Graphical abstract



Highlights

- Combining morphology and biology represents the most useful approach to select candidates for LT in the context of HCC.
- To maintain post-LT 5-year incidence of *"HCC-related death"* <30%, Metroticket 2.0 criteria had to be restricted in some cases.
- Adding mRECIST criteria to Metroticket 2.0 led to correct reclassification of 9.4% of patients with *"HCC-related death"*.
- This additional information can be used to better judge the suitability of candidates for LT following neoadjuvant therapies.

Authors

Alessandro Cucchetti, Matteo Serenari, Carlo Sposito, …, Giorgio Ercolani, Antonio Daniele Pinna, Matteo Cescon

Correspondence

aleqko@libero.it (A. Cucchetti).

Lay summary

In the context of liver transplantation for patients with hepatocellular carcinoma, prediction models are used to ensure that the risk of recurrence after transplantation is acceptably low. The Metroticket 2.0 model has been proposed as an accurate predictor of *"tumour-related death"* after liver transplantation. In the present study, we show that its accuracy can be improved by incorporating information relating to the radiological responses of patients to neoadjuvant therapies.

https://doi.org/10.1016/j.jhep.2020.03.018

Research Article
Liver Transplantation

JOURNAL OF HEPATOLOGY

Including mRECIST in the Metroticket 2.0 criteria improves prediction of hepatocellular carcinoma-related death after liver transplant

Alessandro Cucchetti[1,2,*], Matteo Serenari[1,3], Carlo Sposito[4,5], Stefano Di Sandro[6], Cristina Mosconi[3], Ilaria Vicentin[5,7], Enrico Garanzini[8], Vincenzo Mazzaferro[4,5], Luciano De Carlis[6,9], Rita Golfieri[3,10], Carlo Spreafico[8], Angelo Vanzulli[5,7], Vincenzo Buscemi[6], Matteo Ravaioli[1,3], Giorgio Ercolani[1,2], Antonio Daniele Pinna[1,3], Matteo Cescon[1,3]

[1]Department of Medical and Surgical Sciences – DIMEC, Alma Mater Studiorum - University of Bologna, Bologna, Italy; [2]Morgagni – Pierantoni Hospital, Forlì, Italy; [3]S.Orsola – Malpighi Hospital, Bologna, Italy; [4]General Surgery and Liver Transplantation Unit, Fondazione IRCCS Istituto Nazionale Tumori di Milano, Milan, Italy; Università degli Studi di Milano, Milan, Italy; [5]Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy; [6]General Surgery and Abdominal Transplantation Unit, ASST Niguarda Hospital, Milan, Italy; [7]Department of Diagnostic and Interventional Radiology, ASST Niguarda Hospital, Milan, Italy; [8]Department of Radiology, Fondazione IRCCS Istituto Nazionale Tumori di Milano, Milan, Italy; [9]Department of Medicine and Surgery, University of Milano-Bicocca, Italy; [10]Department of Specialized, Diagnostic and Experimental Medicine – DIMES, Alma Mater Studiorum - University of Bologna, Bologna, Italy

Background & Aims: In the context of liver transplantation (LT) for hepatocellular carcinoma (HCC), prediction models are used to ensure that the risk of post-LT recurrence is acceptably low. However, the weighting that 'response to neoadjuvant therapies' should have in such models remains unclear. Herein, we aimed to incorporate radiological response into the Metroticket 2.0 model for post-LT prediction of *"HCC-related death"*, to improve its clinical utility.

Methods: Data from 859 transplanted patients (2000-2015) who received neoadjuvant therapies were included. The last radiological assessment before LT was reviewed according to the modified RECIST criteria. Competing-risk analysis was applied. The added value of including radiological response into the Metroticket 2.0 was explored through category-based net reclassification improvement (NRI) analysis.

Results: At last radiological assessment prior to LT, complete response (CR) was diagnosed in 41.3%, partial response/stable disease (PR/SD) in 24.9% and progressive disease (PD) in 33.8% of patients. The 5-year rates of *"HCC-related death"* were 3.1%, 9.6% and 13.4% in those with CR, PR/SD, or PD, respectively (p <0.001). Log_{10}AFP (p <0.001) and the sum of number and diameter of the tumour/s (p <0.05) were determinants of *"HCC-related death"* for PR/SD and PD patients. To maintain the post-LT 5-year incidence of *"HCC-related death"* <30%, the Metroticket 2.0 criteria were restricted in some cases of PR/SD and in all cases with PD, correctly reclassifying 9.4% of patients with *"HCC-related death"*, at the expense of 3.5% of patients who did not have the event. The overall/net NRI was 5.8.

Conclusion: Incorporating the modified RECIST criteria into the Metroticket 2.0 framework can improve its predictive ability. The additional information provided can be used to better judge the suitability of candidates for LT following neoadjuvant therapies.

Lay summary: In the context of liver transplantation for patients with hepatocellular carcinoma, prediction models are used to ensure that the risk of recurrence after transplantation is acceptably low. The Metroticket 2.0 model has been proposed as an accurate predictor of *"tumour-related death"* after liver transplantation. In the present study, we show that its accuracy can be improved by incorporating information relating to the radiological responses of patients to neoadjuvant therapies.

Introduction

Hepatocellular carcinoma (HCC) represents one of the leading indications for liver transplantation (LT).[1] The main obstacle to applying LT to all patients with HCC is the need to minimize the post-operative recurrence of the tumour. Several studies have shown how radiological features and biological surrogates, such as alpha-fetoprotein (AFP) and response to pre-LT neoadjuvant therapies, can provide important information on candidate selection, prioritization and, ultimately, survival after LT.[2–11] The goal should be to incorporate this additional information into refined transplantability criteria.

Within the prognostic indexes proposed in the modern clinical scenario, the Metroticket 2.0 criteria were developed under a competing-risk framework. Such an approach overcomes a fundamental bias when analysing post-LT outcomes, especially in the presence of highly effective antivirals for HCV.[12–14] In fact, the primary outcome of interest in the present setting should be death due to HCC recurrence, whereas deaths due to other causes, such as hepatitis recurrence, are competing events which could mask the real prognostic impact of pre-transplant radiological and biological features on HCC recurrence and

Keywords: Hepatocellular carcinoma; Liver transplant; Neoadjuvant therapy; Survival; mRECIST.
Received 11 July 2019; received in revised form 21 February 2020; accepted 10 March 2020; available online 20 March 2020
* Corresponding author. Address: Department of Medical and Surgical Sciences – DIMEC, Alma Mater Studiorum - University of Bologna, Via Albertoni 15, 40138 Bologna, Italy.
E-mail address: aleqko@libero.it (A. Cucchetti).
https://doi.org/10.1016/j.jhep.2020.03.018

tumour-related survival.[3,14] However, the weight of the response to neoadjuvant therapies remained a partially unsolved question in these criteria model, as well as in other criteria currently proposed, such as the AFP-French model or the HALT-HCC.[2,4] On one hand, Metroticket 2.0 showed that the radiological behaviour prior to LT correlated with pre-transplant tumour size, number and AFP, which formed the final prognostic model, but on the other hand it did not specify whether the final tumour burden was the consequence of a partial response to neoadjuvant therapies, or of a stable or progressive disease (PD) of an untreated tumour or after neoadjuvant treatments. To be unequivocal, a single vital T2 tumour with moderate AFP levels, resulting from a down-staging protocol, has a different biology to a similar tumour that is stable after neoadjuvant therapy or results from PD. In the first, we can perceive a favourable biology, whereas, in the latter case, we can perceive an unfavourable aggressiveness with negative post-transplant consequences.

Thus, the aim of the present study was to re-evaluate the relationship between tumour features and radiological behaviour prior to LT, to refine the Metroticket 2.0 criteria by incorporating the *"biological history"* of transplanted tumours.

Materials and methods

Part of the cohort of the Metroticket 2.0 model was reviewed and updated, the present cohort included only cases which formed the derivation dataset. After update, the study population started with 1,137 patients with known HCC who underwent LT for HCC between January 2000 and December 2015 at 3 tertiary referral hepato-biliary and transplant centres in Italy. Incidental HCCs were not included in the present data collection. Priority policy and transplantation criteria were those already published in the Metroticket 2.0.[3] From the starting population, 149 patients were excluded because of incomplete dataset, and/or because of the absence of pre-LT radiological evaluation, resulting in a population of 988 patients. Only patients who received neoadjuvant treatments were retained, leading to the exclusion of 129 untreated patients. Consequently, the final study population consisted of 859 patients. The use of present data was approved by the National Cancer Institute Institutional Review Board and fulfilled the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data.

Neoadjuvant therapies

For all patients, the history of neoadjuvant therapies, was collected. Oncological data were obtained at referral to the tumour board meetings and after their application prior to LT using the last available imaging data. The decision for any eventual specific neoadjuvant therapy was decided during such multidisciplinary groups based on tumour stage and liver function. The adopted treatments did not fully accomplish the international guideline suggestions but were individualized based on patient and tumour features, as well as the perceived probability of success. Hepatic resection was offered to HCC patients deemed resectable in the setting of a liver transplant program, considering both "salvage" and "bridge" perspectives.

Radiological assessment

At radiology, only nodules with a maximum diameter of at least 1 cm counted as HCC nodules, providing that the international criteria released during the study period were fulfilled. The radiological assessment before LT was accomplished according to the modified RECIST (mRECIST) criteria.[15] The mRECIST data were prospectively collected, apart from in patients enlisted before 2010, in whom they were retrospectively reviewed. In addition, the maximum diameter of the active nodule/s was recorded as in RECIST 1.1[16] for use in the Metroticket 2.0 model. Candidates were followed during waiting list within a 3-month radiological surveillance which included contrast-enhanced ultrasound (CEUS) with a second-generation ultrasonography contrast agent (SonoVue, Bracco Imaging S.p.A., Milano, Italy), tri- or quadriphasic CT scan or MR imaging, depending on clinical indications and suitability. Complete radiological response (CR) was diagnosed when imaging showed the absence of any intratumoural enhancement during arterial phase.[15]

The inter-observer (inter-centre) agreement was explored through the cross-check of mRECIST of 180 random cases between the 3 centres, to provide an estimation of the bias related to this issue. This sample was representative of the entire study population with a 95% of confidence and a 10% margin of error. The final mRECIST evaluation considered for all the analyses was that of the corresponding centre which transplanted the patient.

Pathological assessment

At histology, the explanted livers were sampled to evaluate all macroscopically suspect nodules, and neoplastic masses were thoroughly examined. The complete absence of any residual vital tumour tissue defined the complete pathological response (CPR). When tumour tissue was detected, the number and size of active HCC nodules was reported together with the presence or absence of microscopic vascular invasion. In the present study, tumour grading was not included because of too much missing data. In cases where LT was performed early (<90 days) after hepatic resection due to post-operative liver failure, the pathological findings of the resected specimen and the explanted liver were considered altogether.

Metroticket 2.0 framework

Patients with CR after neoadjuvant treatments were considered as having 0 nodules with 0 diameter at pre-LT assessment. As already reported, the sum of number and size of the largest tumour (*i.e.* 1 nodule of 6 cm = up-to-7) was considered in the subsequent competing-risk regression.[3] The Metroticket 2.0 criteria were constructed to maintain the 5-year *"HCC-related death"* incidence below 30% to allow for confidence bands. This safety threshold was fulfilled in 3 pre-LT radiological clinical situations, defining the Metroticket 2.0 criteria IN:

1. HCC at pre-LT radiology within the up-to-7 criteria, if AFP <200 ng/ml;
2. HCC at pre-LT radiology within the up-to-5 criteria, if AFP 200–400 ng/ml;
3. HCC at pre-LT radiology within the up-to-4 criteria, if AFP 400–1000 ng/ml;

Follow-up and outcome definitions

All survival analyses fulfilled the competing-risk concept. Date of death/last censoring, date of diagnosis of HCC recurrence and cause of death were collected. For recurrence-free survival, death without evidence of HCC recurrence was considered as a competing event. Patient survival was the primary outcome measure where death was defined as *"HCC-related"* when

consequent to a documented HCC recurrence, either intra- or extrahepatic, submitted or not to additional non-transplant therapies, which unequivocally lead to a progressive worsening of performance status until death. Death because of tumour recurrence was also defined when liver function worsened as a consequence of intra-hepatic tumour spread, such as massive liver involvement or development of tumoural macroscopic vascular invasion. All the other causes of death were defined as *"non-tumour-related"* and considered as competing events.

Statistical analysis

No missing data were present. Continuous variables were reported as medians and IQRs. The Kruskal-Wallis and the Fisher exact tests were adopted as appropriate. Survival probabilities were estimated using the Kaplan-Meier method. The last censoring was on January 2019. Competing-risk analyses were performed using the methodologies provided by Scrucca and Fine & Gray[17,18] and repeated for each mRECIST criteria considered. Such stratification was adopted to overcome biases generated by collinearity between Metroticket 2.0 criteria and mRECIST criteria.[15,16] The added value of incorporating radiological response into the Metroticket 2.0 criteria was reported through the category-based net reclassification improvement (NRI),[19] defining *"HCC-related death"* within 5 years from LT as events, and patients alive at 5 years as non-events. Analyses were performed using STATA and R-Project. Finally, mRECIST agreement between centres on the sample of 180 random cases was assessed using k-Cohen.

Results

The baseline features of 859 patients are summarized in Table 1. Median follow-up was 5 years (IQR 3.0–8.1). During this time, 111 patients had post-transplant HCC recurrence and 208 deaths were registered, of which 81 were tumour related. Recurrence and "*HCC-related death*" rates are detailed in Table 1.

The median time from last radiological assessment and LT was of 2.2 months (IQR 1.0–4.1). The last radiological assessment was performed with CT in 636 patients (74.0%), with MR in 197 (23.0%) and with CEUS in 26 patients (3.0%). CR was diagnosed in 41.3%, partial response/stable disease (PR/SD) in 24.9%, and PD in 33.8%.

Competing-risk analysis (Fig. 1A) showed that patients with CR had 1-, 3- and 5-year *"HCC-related death"* rates of 0.3%, 1.8% and 3.1%, respectively; patients with PR/SD had rates of 1.9%, 6.1% and 9.6% and patients with PD had rates of 2.8%, 10.9% and 13.4%, respectively (p <0.001).

Complete radiological response

The 355 patients with CR at last radiological assessment prior to LT, received a total of 523 neoadjuvant treatments (Table S1 and S2). The median time from last radiological assessment to LT was 2.3 months (IQR 1.1–4.4). At final histology, a CPR was demonstrated in 46.5% of cases. As depicted in Fig. 1B, recipients with CPR had a 1-, 3- and 5-year rates of *"HCC-related death"* of 0%, 0.6% and 1.4%, respectively (corresponding to 2 patients who developed bone metastases within 18 months from LT). In contrast, recipients without CPR had a 1-, 3- and 5-year rates of *"HCC-related death"* of 0.5%, 2.7% and 4.4%, respectively (p = 0.006). The pre-LT AFP was not found related to this survival endpoint (Table 2).

Table 1. Clinical features of recipients who received neoadjuvant therapies.

Clinical features	n = 859
Age [years (median; IQR)]	57 (52–62)
Male, n (%)	754 (87.8)
Hepatitis C, n (%)	489 (56.9)
Hepatitis B, n (%)	275 (32.0)
Alcohol, n (%)	167 (19.4)
Other, n (%)	35 (4.1)
Within Milan criteria at diagnosis/before listing (%)	751 (87.4)
Number of neoadjuvant therapies (median; IQR)*	1 (1–2)
Number of therapies ≥3, n (%)	104 (12.1)
Time from diagnosis to LT [months (median; IQR)]	13.1 (8.0–22.7)
Time spent in waiting list [months (median; IQR)]	4.3 (2.2–10.0)
>6 months, n (%)	341 (39.7)
Time from last radiological assessment and LT [months (median; IQR)]	2.2 (1.0–4.1)
Complete Response, n (%)	355 (41.3)
Partial Response/Stable Disease, n (%)	214 (24.9)
Progressive disease, n (%)†	290 (33.8)
Number of vital tumours	
None, n (%)	355 (41.3)
Single nodule, n (%)	258 (33.6)
2-3 nodules, n (%)	207 (24.1)
>3 nodules, n (%)	39 (4.5)
Largest vital tumour at pre-LT radiology [cm, (median; IQR)]†	2.1 (1.5–3.0)
Within Milan criteria at pre-LT radiology (%)	803 (93.5)
Last AFP before LT [ng/ml (median; IQR)]	8.1 (4.0–22)
Log_{10} AFP [ng/ml (median; IQR)]	0.91 (0.60–1.34)
Complete pathological response at histology, n (%)	214 (24.9)
Recurrence-rate, %‡	
1-year (95% CI)	4.0 (2.8–5.5)
3-year (95% CI)	10.4 (8.4–12.6)
5-year (95% CI)	12.5 (10.3–14.9)
HCC-related death, %‡	
1-year (95% CI)	1.5 (0.8–2.6)
3-year (95% CI)	5.9 (4.4–7.7)
5-year (95% CI)	8.1 (6.4–10.2)

FP, alpha-fetoprotein; HCC, hepatocellular carcinoma; LT, liver transplantation.
*Refer to Supplementary table for details.
†Diameters of non-vital tumours were excluded from the calculation.
‡Calculated through competing-risk analysis, with death without recurrence/for other causes as competing events.

Partial response/stable disease

The 214 patients with PR/SD at last radiological assessment received a total of 326 neoadjuvant therapies (Table S1 and S2). The median time from last radiological assessment and LT was of 2.0 months (IQR 1.0–3.9). At final histology, the absence of HCC was observed in 9.4% of cases. Competing-risk regression (Table 2) showed Log_{10}AFP (p <0.001) and the sum of number and diameter of the tumour/s (p <0.001) as determinants of *"HCC-related death"*.

Progressive disease

The 290 patients showing PD prior to LT received a total of 479 neoadjuvant therapies (Table S1 and S2). The median time from last radiological assessment and LT was 2.2 months (IQR 0.9–4.0). At final histology, the absence of HCC was observed in 10.0% of cases. Competing-risk regression (Table 2) confirmed Log_{10}AFP (p <0.001) and the sum of number and diameter of the tumour/s (p = 0.004) as determinants of *"HCC-related death"* in these patients.

Radiological response within the Metroticket 2.0 framework

The original Metroticket 2.0 criteria were reviewed based on the hazard ratios of each sub-group. To accomplish a cumulative

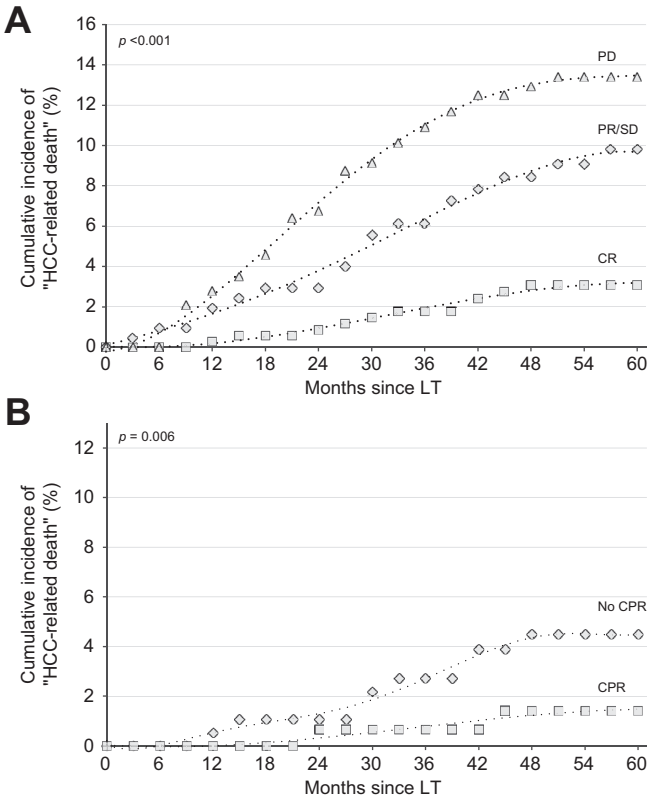


Fig. 1. Cumulative incidences of *"HCC-related death"*. (A) Stratified by last mRECIST assessment and by (B) absence/presence of complete pathological response at final histology in 355 patients showing complete radiological response at last pre-LT imaging. Error bars represent standard errors. *p* values derived from log-rank test for competing-risk analyses.[18] CPR, complete pathological response; CR, complete response; HCC, hepatocellular carcinoma; LT, liver transplantation; mRECIST, modified RECIST; PD, progressive disease; PR/SD, partial response/stable disease.

incidence of *"HCC-related death"* of <30% 5 years after transplant, the criteria were adjusted as reported in Fig. 2. Dimensional criteria were restricted in some cases of PR/SD and in all cases of PD prior to LT. Consequently, 35 patients within Metroticket 2.0 criteria did not fulfil the present radiological-adjusted criteria. These 35 patients had 1-, 3-, and 5-year rates of *"HCC-related death"* of 2.9%, 17.4% and 25.0%, respectively.

As detailed in the Table 3, in comparison to the Metroticket 2.0 criteria, the inclusion of radiological response resulted in a correct reclassification of 9.4% of patients who died from *"HCC-related death"* within 5 years from LT but at the expense of 3.5% of patients who did not have the event, however, the overall NRI was positive at 5.8. In comparison to the Milan criteria, the application of radiological response determined a correct reclassification of 10.9% of patients who died from *"HCC-related death"* within 5 years from LT and of 2.0% of patients who did not have the event, with an overall NRI of 12.9.

Inter-centre agreement on radiological response

In the sample of 180 radiological evaluations for the assessment of mRECIST agreement, the proportion of concordant pairs was 77.8%, with a kappa of 0.67 (95% CI 0.55–0.79) indicating substantial agreement between centres (Table 4). When PR and SD were pooled, the proportion of concordant pairs increased to 84.4% with a kappa of 0.75 (95% CI 0.62–0.87).

Discussion

Assessing LT eligibility in patients with HCC is an evolving field of research.[20,21] The need for accurate outcome prediction is crucial to ensure that the waiting list includes those candidates who will benefit most, while excluding those with an unfavourable prognosis.[22,23] The present study fulfils the efforts of identifying *"transplantable tumours"* as accurately as possible through the estimation of the impact of response to neoadjuvant therapies. Recent studies focused on the combination of tumour features and biological markers, and, on this background, the AFP-French model, the HALT-HCC and the Metroticket 2.0 models were the most reliable in the clinical setting.[2–4] These models shared common features but were each unique. The AFP-French model was based on tumour number/size and AFP at listing[2] but no data on response to neoadjuvant therapy were reported or considered. In the HALT-HCC model the history of neoadjuvant therapy was recorded (59.6% of the derivation cohort received locoregional therapy) but the result was not significant for overall survival prediction in the final regression model.[4] The Metroticket 2.0 depicted the relationships between the radiological behaviour through RECIST 1.1 criteria and tumour features at transplant, using only the latter for HCC-specific post-LT survival.[3] Present results suggest that the inclusion of mRECIST criteria within the Metroticket 2.0 framework can provide useful information to refine outcome prediction.

In 2018, the Metroticket 2.0 calculator was developed with the aim of predicting post-LT HCC-specific survival.[3] This survival endpoint was introduced to appropriately weigh the oncologic determinants of post-LT death in the presence of other confounding causes of death. This calculator showed good accuracy on an Eastern series of patients and was generalizable to different populations.[3] However, the aim of that study was to provide a prognostic tool that combined both simplicity of use and accuracy, and for this reason, at that stage, the variables taken into account in the model were only tumour morphology (sum of size and number) and AFP. The relationships between radiological features of treated and untreated HCCs and pre-LT

Table 2. Coefficients resulting from stratified competing-risk analysis on HCC-related death.

mRECIST	Number + diameter	*p* value	Log_{10} AFP	*p* value
CR	n.a.	–	0.172 (−0.748–1.092)	0.714
PR/SD	0.236 (0.106–0.365)	<0.001	0.934 (0.424–1.443)	<0.001
PD	0.188 (0.061–0.315)	0.004	0.870 (0.524–1.216)	<0.001

The baseline cumulative hazard at 5-year for PR/SD was 0.010109 and for PD was 0.016729.
The predicted 5-year probability of "HCC-related death" can be obtained solving the following equation: pi = [baseline cumulative hazard x exp (coefficient x number + diameter + coefficient x Log_{10}AFP)] and 5-year = 1-exp(-pi).
The median AFP in patients with CR was 6.8 ng/ml (IQR 3.1–17.6), in patients with PR/SD was 8.6 (IQR 4.0–22.0) and in patients with PD was 10.0 (IQR 5.0–32.0).
AFP, alpha-fetoprotein; CR, complete response; HCC, hepatocellular carcinoma; mRECIST, modified RECIST; n.a., not applicable; PD, progressive disease; PR/SD, partial response/stable disease.

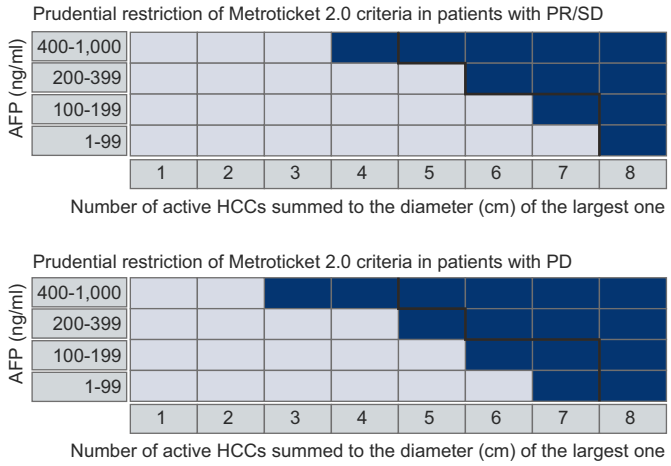


Fig. 2. Careful restriction of Metroticket 2.0 criteria for transplantable HCC. Based on morphology (largest vital tumour size + number of vital tumour nodules) and biology (AFP and last radiological mRECIST assessment) after neoadjuvant therapies. The darkest areas describe a 5-year *"HCC-related death"* above the safety threshold of 30%. The black bold line represents the original Metroticket 2.0 criteria. AFP, alpha-fetoprotein; HCC, hepatocellular carcinoma; mRECIST, modified RECIST; PD, progressive disease; PR/SD, partial response/stable disease.

tumour number and diameter were described, but not included in the calculator because of their correlations.[3] Nevertheless, neoadjuvant locoregional therapies are commonly applied in

Table 3. NRI of inclusion of mRECIST evaluation into the Metroticket 2.0 model in comparison to original Metroticket 2.0 and Milan criteria.

Criteria		Metroticket + mRECIST		
Events*		IN	OUT	Total
Metroticket 2.0	IN	40	6	46
	OUT	0	18	18
	Total	40	24	64
Non-events		IN	OUT	Total
Metroticket 2.0	IN	368	15	383
	OUT	1	11	12
	Total	369	26	395
Events correctly reclassified by mRECIST assessment *vs.* Metroticket 2.0 Higher = 6/64 = 9.4%; Lower = 0/64 = 0%; NRI events = 9.4 – 0 = 9.4% Non - events correctly reclassified by mRECIST assessment *vs.* Metroticket 2.0 Higher = 15/395= 3.8%; Lower = 1/395 = 0.3%; NRI events = 0.3 – 3.8 = −3.5% Overall NRI = 9.4 – 3.5 = 5.8				

Criteria		Metroticket + mRECIST		
Events*		IN	OUT	Total
Milan criteria	IN	39	8	47
	OUT	1	16	17
	Total	40	24	64
Non-events		IN	OUT	Total
Milan criteria	IN	356	5	361
	OUT	13	21	34
	Total	369	26	395
Events correctly reclassified by mRECIST + Metroticket *vs.* Milan criteria Higher = 8/64 = 12.5%; Lower = 1/64 = 1.6%; NRI events = 12.5 – 1.6 = 10.9% Non - events correctly reclassified by mRECIST + Metroticket *vs.* Milan criteria Higher = 5/395= 1.3%; Lower = 13/395 = 3.3%; NRI events = 3.3 – 1.3 = 2.0% Overall NRI = 10.9 + 2.0 = 12.9				

NRI of Metroticket 2.0 *vs.* Milan criteria is 12.9 – 5.8 = 7.1.
HCC, hepatocellular carcinoma; LT, liver transplantation; mRECIST, modified RECIST; NRI, net reclassification improvement.
*NRI is calculated among patients who died for HCC recurrence within 5 years from LT (events = 64) and those who were alive without recurrence at this temporal endpoint (non-events = 395).

Table 4. Overall distribution of mRECIST evaluation pairs from the 3 participating centres.

	CR	PR	SD	PD
CR	35	4	1	0
PR	3	21	2	0
SD	2	4	3	0
PD	0	1	3	11

The number of observed agreements was 70/90 pairs (77.8%). The kappa was 0.67 (95% CI 0.55–0.79) indicating substantial agreement between centres (k = 0.00-0.20: slight agreement; 0.21-0.40: fair agreement; 0.41-0.60: moderate agreement; 0.61-0.80: substantial agreement; 0.81-1.00: almost perfect agreement.
Numbers represent the sum of pairs of mRECIST evaluations between centres #1 *vs.* #2, #2 *vs.* #3 and #1 *vs.* #3.
CR, complete response; mRECIST, modified RECIST; PD, progressive disease; PR/SD, partial response/stable disease.

HCC candidates, as bridging or down-staging procedures, and the response to these treatments can suggest a more, or less, aggressive disease from a biological point of view, which deserves consideration.[5–7] Lee *et al.* recently showed that the incidence of post-LT recurrence, analysed through a competing-risk approach, was significantly worse in patients who did not respond to pre-LT neoadjuvant therapies, being up to about 26% at 5 years.[5] On the contrary, the same figure was <10% in cases showing complete or partial response according to mRECIST criteria.[5] Two analyses of the US Multicenter HCC Transplant Consortium, among more than 3,000 patients receiving or not receiving pre-LT locoregional therapies, highlighted that the response to neoadjuvant treatment and unfavourable wait-list AFP were the main determinants of post-LT recurrence, resulting in a hazard ratio of ~2.1.[6,7] In the present study, all this information was embedded within the Metroticket 2.0 criteria through the inclusion of mRECIST, in order to provide some additional clinically useful information.

First, we observed that pre-LT radiological assessment provided a satisfactory discrimination of *"HCC-related death"* 5 years after LT, with rates of 3.1%, 9.6% and 13.4% for patients showing CR, PR/SD and PD, respectively, at the time of transplant. These data confirm that progression on the waiting list represents the clinical expression of a more aggressive tumour behaviour, that in a non-negligible percentage of cases might turn into post-LT death because of recurrence.[5–7,24] However, the observation of PD prior to LT should not be considered as a contraindication to LT, it simply calls for prudence in the selection of these patients by narrowing the dimensional criteria. This restriction would allow patients with PD to obtain an outcome not far from that of patients with PR/SD and with larger tumour burden. On the other hand, patients showing CR because of neoadjuvant therapies at the time of LT had a very low risk of post-LT *"HCC-related death"*, and may be considered as *"ideal candidates"*, providing that the initial HCC stage was consistent with an indication to LT. Additionally, patients showing CR also represented a category for which it is not necessary to provide any particular priority on the waiting list.[22,23,25]

Determinants of HCC-specific survival were the usual, as shown in Table 2. However, stratification for radiological response, necessary to minimize possible collinearity between mRECIST and dimensional assessments in the original Metroticket 2.0 criteria, led to some interesting results. First, AFP did not turn out to be related to *"HCC-related death"* for patients showing CR to neoadjuvant therapies. This result was almost expected considering that a negligible proportion of these patients had AFP >20 ng/ml. Conversely, both tumour size +

number and AFP turned out to be consistently associated with *"HCC-related death"* in patients with PR/SD and PD. The absolute effect of size + number and AFP was similar among PR/SD patients and PD patients, as suggested by similar coefficients, but their relative impact on outcomes increased with the worsening of response to neoadjuvant therapies, as suggested by increasing baseline cumulative hazard (for PR/SD was about 0.0101 and for PD was 0.0167). Overall, when analysed in the context of the AFP-adjusted criteria proposed in the Metroticket 2.0, the present results confirmed the validity of the original criteria for most patients who get to LT with a PR/SD, but also suggest restricting the indication for some patients with PR/SD and for all patients showing PD.

The end-result of the adoption of radiological response assessment was that an additional 9.4% of patients who died because of *"HCC-related death"* within 5 years from LT would have been identified compared to using the Metroticket 2.0 application alone. This increased sensitivity would be paid for by a decrease in specificity of 3.5%, but the positive overall NRI of 5.8 suggests that the benefits outweigh the harms of such an approach, providing clinical utility. A brief explanation and comparison of this value is necessary at this point. The NRI value derives from considering its 2 reported components (events and non-events), an identical point estimate of this statistic may have different interpretations depending on its components.[19,26] The present NRI value showed that this was positive, thus, that the new classification had positive effects. As a comparator, the AFP-French model showed in its training cohort a NRI of 11.4 over Milan criteria, as the end-result of a correct reclassification of 6.9% of patients who had HCC recurrence within 5 years from LT, at the expense of 4.5% of patients who did not have the event. The current model showed an NRI of 12.9 over Milan criteria as the result of a correct reclassification of 10.9% of events and of 2.0% of non-events. However, higher NRI does not necessarily mean higher accuracy[24] and details about reclassification of events and non-events among risk classes were reported herein to indicate the direction and the magnitude of reclassification.

Another interesting finding of the present study was the observation that radiological diagnosis of CR can be considered quite inaccurate. The median time from last radiological assessment and LT was <3 months in most cases, and radiological assessment was performed by dedicated radiologists at high volume centres. However, in the 355 patients showing radiological absence of vital tumours, CPR was confirmed in only 46.5% of cases. These findings are in line with most prospective and retrospective studies, in which a radiological CR was confirmed by a CPR in 50–82.4% of cases, highly depending on the type and number of pre-LT treatments.[6,9,24,27,28] The accuracy of the diagnosis of complete radiological response acquires importance in the light of post-transplant outcomes, with an incidence of *"HCC-related death"* at 5 years of only 1.4% when CR was confirmed by CPR. Interestingly, the present end-result was identical to that previously reported by Agopian *et al.* in 501 transplanted patients.[29] This low radiological accuracy will probably cause a disagreement between radiological PR/SD/PD and final pathological evaluation of tumour burden. In this sense, pathological assessment would provide the optimal information to estimate the risk of recurrence, as in the original Metroticket published in 2009, but lacks clinical applicability from an intention-to-treat point of view before LT. Consequently, and at present, clinical decisions should still be based on an imperfect system which already includes misclassification probabilities.

This study has some limitations: first, this is a retrospective study in which a part of the radiological evaluation had been re-assessed according to mRECIST because this data was not available before 2010. This aspect, together with a lack of central radiological review, may weaken radiological response evaluation. The large sample size and the homogeneous radiological protocols among the 3 participating centres can counterbalance this potential bias. In fact, as suggested by the FDA Clinical Trial Imaging Endpoint Process Standards 2018, centralized image interpretation processes may provide more verifiable and uniform reader training, but centralized image interpretation is not always critical when using some aspects of quantitative imaging.[30] If the quantitative measure is widely performed and reported in clinical medicine, as for mRECIST evaluation, centralization would have been necessary if radiological centres with little experience in the evaluation of these criteria had been involved in the present study.[30] Nevertheless, the analysis of mRECIST agreement between the participating centres suggested that a source of uncertainty is present, albeit unavoidable in clinical practice. It should be noted however, that when pooling PR/SD cases, as in the present study, the proportion of agreements was up to 84.4%, which represent a clinically useful result.

Second, the assessment of radiological response included CEUS, CT and MR, depending on clinical indications and suitability. Consequently, and especially for patients submitted to transarterial chemoembolization, mRECIST assessment would not be ideal.[27] Since radiological protocols and expertise were similar between participating centres, it is unlikely that a centralized image review would increase accuracy in the mRECIST assessment, as stated before. Present data simply represent the end-results of the common clinical practice where it would be unlikely to follow all patients with MR. Finally, radiological evaluation and post-LT outcomes were not analysed according to the type of locoregional or surgical treatment applied. This choice was made upfront, because the selection of treatments in the pre-LT period is a dynamic process that depends on several variables that might change over this period, and that cannot be fully considered in a retrospective study. The scope of the present study that was to provide a static pre-LT selection criterion that might guide clinicians in the decision process when a donor is available, namely AFP and last radiological evaluation in terms of mRECIST criteria and morphological tumour characteristics.

In conclusion, we have gone a step further in the selection process for LT, highlighting that identifiable patients with PR/SD or PD should be carefully considered based on morphological/biological criteria at the time of LT, to maintain an acceptably low risk of post-LT *"HCC-related death"*.

Abbreviations

AFP, alpha-fetoprotein; CEUS, contrast-enhanced ultrasound; CPR, complete pathological response; CR, complete response; HCC, hepatocellular carcinoma; LT, liver transplantation; NRI, net reclassification improvement; PD, progressive disease; PR, partial response; SD, stable disease.

Financial support

The authors received no financial support to produce this manuscript.

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

A. Cucchetti and M. Serenari had the original idea; M. Serenari, C. Sposito, S. Di Sandro, V. Buscemi and M. Ravaioli collected clinical data; R. Golfieri, C. Spreafico, A. Vanzulli collected and reviewed radiological data; V. Mazzaferro, L. De Carlis, A.D. Pinna, M.. Cescon and G. Ercolani performed most of liver transplants, A. Cucchetti performed all the analyses, C. Sposito and M. Serenari helped in wrote the manuscript; C. Mosconi, E. Garanzin and I. Vicentin provided inter-centre agreement.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2020.03.018.

References

Author names in bold designate shared co-first authorship

[1] Kim WR, Lake JR, Smith JM, Schladt DP, Skeans MA, Harper AM, et al. OPTN/SRTR 2016 Annual Data Report: Liver. Am J Transplant 2018;18(Suppl 1):172–253.

[2] Duvoux C, Roudot-Thoraval F, Decaens T, Pessione F, Badran H, Piardi T, et al. Liver transplantation for hepatocellular carcinoma: a model including alpha-fetoprotein improves the performance of Milan criteria. Gastroenterology 2012;143:986–994.

[3] Mazzaferro V, Sposito C, Zhou J, Pinna AD, De Carlis L, Fan J, et al. Metroticket 2.0 model for analysis of competing risks of death after liver transplantation for hepatocellular carcinoma. Gastroenterology 2018;154:128–139.

[4] Sasaki K, Firl DJ, Hashimoto K, Fujiki M, Diago-Uso T, Quintini C, et al. Development and validation of the HALT-HCC score to predict mortality in liver transplant recipients with hepatocellular carcinoma: a retrospective cohort analysis. Lancet Gastroenterol Hepatol 2017;2:595–603.

[5] Lee DD, Samoylova M, Mehta N, Musto KR, Roberts JP, Yao FY, et al. The mRECIST classification provides insight into tumor biology for patients with hepatocellular carcinoma awaiting liver transplantation. Liver Transpl 2019;25:228–241.

[6] DiNorcia J, Florman SS, Haydel B, Tabrizian P, Ruiz RM, Klintmalm GB, et al. Pathologic response to pretransplant locoregional therapy is predictive of patient outcome after liver transplantation for hepatocellular carcinoma: analysis from the US multicenter HCC transplant consortium. Ann Surg 2020;271(4):616–624.

[7] Agopian VG, Harlander-Locke MP, Ruiz RM, Klintmalm GB, Senguttuvan S, Florman SS, et al. Impact of pretransplant bridging locoregional therapy for patients with hepatocellular carcinoma within milan criteria undergoing liver transplantation: analysis of 3601 patients from the US multicenter HCC transplant consortium. Ann Surg 2017;266:525–535.

[8] Vitale A, Volk ML, De Feo TM, Burra P, Frigo AC, Ramirez Morales R, et al. A method for establishing allocation equity among patients with and without hepatocellular carcinoma on a common liver transplant waiting list. J Hepatol 2014;60:290–297.

[9] Toso C, Meeberg G, Hernandez-Alejandro R, Dufour JF, Marotta P, Majno P, et al. Total tumor volume and alpha-fetoprotein for selection of transplant candidates with hepatocellular carcinoma: a prospective validation. Hepatology 2015;62:158–165.

[10] Halazun KJ, Tabrizian P, Najjar M, Florman S, Schwartz M, Michelassi F, et al. Is it time to abandon the milan criteria?: results of a bicoastal US collaboration to redefine hepatocellular carcinoma liver transplantation selection policies. Ann Surg 2018;268:690–699.

[11] Lai Q, Vitale A, Iesari S, Finkenstedt A, Mennini G, Onali S, et al. The intention-to-treat effect of bridging treatments in the setting of milan criteria-in patients waiting for liver transplantation. Liver Transpl 2019;25:1023–1033.

[12] Belli LS, Perricone G, Adam R, Cortesi PA, Strazzabosco M, Facchetti R, et al. Impact of DAAs on liver transplantation: major effects on the evolution of indications and results. An ELITA study based on the ELTR registry. J Hepatol 2018;69:810–817.

[13] EASL Clinical Practice Guidelines. Management of hepatocellular carcinoma. J Hepatol 2018;69:182–236.

[14] Sposito C, Cucchetti A, Mazzaferro V. Assessing competing risks for death following liver transplantation for hepatocellular carcinoma. Dig Dis Sci 2019;64:1001–1007.

[15] Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis 2010;30:52–60.

[16] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228–247.

[17] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509.

[18] Scrucca L, Santucci A, Aversa F. Regression modeling of competing risk using R: an in depth guide for clinicians. Bone Marrow Transplant 2010;45:1388–1395.

[19] Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–172.

[20] Costentin CE, Bababekov YJ, Zhu AX, Yeh H. Is it time to reconsider the milan criteria for selecting patients with hepatocellular carcinoma for deceased-donor liver transplantation? Hepatology 2019;69:1324–1336.

[21] Mazzaferro V, Battiston C, Sposito C. Pro (with caution): extended oncologic indications in liver transplantation. Liver Transpl 2018;24:98–103.

[22] Di Sandro S, Bagnardi V, Cucchetti A, Lauterio A, De Carlis R, Benuzzi L, et al. From a philosophical framework to a valid prognostic staging system of the new "comprehensive assessment" for transplantable hepatocellular carcinoma. Cancers (Basel) 2019;11(6):741.

[23] Mazzaferro V. Squaring the circle of selection and allocation in liver transplantation for HCC: an adaptive approach. Hepatology 2016;63:1707–1717.

[24] Manzia TM, Lai Q, Iesari S, Perera MTPR, Komuta M, Carvalheiro A, et al. Impact of remnant vital tissue after locoregional treatment and liver transplant in hepatocellular cancer patients, a multicentre cohort study. Transpl Int 2018 Mar 23. https://doi.org/10.1111/tri.13153 [Epub ahead of print].

[25] Cillo U, Burra P, Mazzaferro V, Belli L, Pinna AD, Spada M, et al. A multistep, consensus-based approach to organ allocation in liver transplantation: toward a "blended principle model". Am J Transplant 2015;15:2552–2561.

[26] Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med 2014;160:122–131.

[27] Chapiro J, Wood LD, Lin M, Duran R, Cornish T, Lesage D, et al. Radiologic-pathologic analysis of contrast-enhanced and diffusion-weighted MR imaging in patients with HCC after TACE: diagnostic accuracy of 3D quantitative image analysis. Radiology 2014;273:746–758.

[28] **Serra C**, **Cucchetti A**, Felicani C, Mosconi C, De Cinque A, Golfieri R, et al. Assessment of radiofrequency ablation efficacy for hepatocellular carcinoma by histology and pretransplant radiology. Liver Transpl 2019;25:88–97.

[29] Agopian VG, Morshedi MM, McWilliams J, Harlander-Locke MP, Markovic D, Zarrinpar A, et al. Complete pathologic response to pretransplant locoregional therapy for hepatocellular carcinoma defines cancer cure after liver transplantation: analysis of 501 consecutively treated patients. Ann Surg 2015;262:536–545.

[30] Available at: https://www.fda.gov/media/81172/download. U.S. Food & Drug Administration. Accessed 17 October 2019.