

The Intention-to-Treat Effect of Bridging Treatments in the Setting of Milan Criteria—In Patients Waiting for Liver Transplantation

Quirino Lai ^{1,2}, Alessandro Vitale,³ Samuele Iesari,^{1,4} Armin Finkenstedt,⁵ Gianluca Mennini,² Simona Onali,⁶ Maria Hoppe-Lotichius,⁷ Tommaso M. Manzia ⁸, Daniele Nicolini,⁹ Alfonso W. Avolio,¹⁰ Anna Mrzljak,¹¹ Branislav Kocman,¹¹ Salvatore Agnes,¹⁰ Marco Vivarelli,⁹ Giuseppe Tisone,⁸ Gerd Otto,⁷ Emmanuel Tsochatzis ⁶, Massimo Rossi,² Andre Viveiros,⁵ Olga Ciccarelli,¹ Umberto Cillo,³ and Jan Lerut¹, for the European Hepatocellular Cancer Liver Transplant Study Group

¹Starzl Abdominal Transplant Unit, Pôle de Chirurgie Expérimentale et Transplantation, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain, Brussels, Belgium; ²Department of General Surgery and Organ Transplantation, Umberto I Hospital, Sapienza University, Rome, Italy; ³Department of Surgery, Oncology, and Gastroenterology, University of Padua, Padua, Italy; ⁴Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy; ⁵Department of Medicine I, Medical University Innsbruck, Innsbruck, Austria; ⁶University College London Institute for Liver and Digestive Health and Royal Free Sheila Sherlock Liver Centre, Royal Free Hospital and University College London, London, United Kingdom; ⁷Department of Transplantation and Hepatobiliary Surgery, University of Mainz, Mainz, Germany; ⁸Department of Transplant Surgery, Polyclinic Tor Vergata Foundation, Tor Vergata University, Rome, Italy; ⁹Unit of Hepatobiliary Surgery and Transplantation, Azienda Ospedaliero-Universitaria Ospedali Riuniti, Polytechnic University of Marche, Torrette Ancona, Italy; ¹⁰Liver Unit, Department of Surgery, Agostino Gemelli Hospital, Catholic University of the Sacred Heart, Rome, Italy; and ¹¹Liver Transplant Centre, Merkur University, Zagreb, Croatia

In patients with hepatocellular carcinoma (HCC) meeting the Milan criteria (MC), the benefit of locoregional therapies (LRTs) in the context of liver transplantation (LT) is still debated. Initial biases in the selection between treated and untreated patients have yielded conflicting reported results. The study aimed to identify, using a competing risk analysis, risk factors for HCC-dependent LT failure, defined as pretransplant tumor-related delisting or posttransplant recurrence. The study was registered at www.clinicaltrials.gov (identification number NCT03723304). In order to offset the initial limitations of the investigated population, an inverse probability of treatment weighting (IPTW) analysis was used: 1083 MC-in patients (no LRT = 182; LRT = 901) were balanced using 8 variables: age, sex, Model for End-Stage Liver Disease (MELD) value, hepatitis C virus status, hepatitis B virus status, largest lesion diameter, number of nodules, and alpha-fetoprotein (AFP). All the covariates were available at the first referral. After the IPTW, a pseudo-population of 2019 patients listed for LT was analyzed, comparing 2 homogeneous groups of untreated ($n = 1077$) and LRT-treated ($n = 942$) patients. Tumor progression after LRT was the most important independent risk factor for HCC-dependent failure (subhazard ratio [SHR], 5.62; $P < 0.001$). Other independent risk factors were major tumor diameter, AFP, MELD, patient age, male sex, and period of wait-list registration. One single LRT was protective compared with no treatment (SHR, 0.51; $P < 0.001$). The positive effect was still observed when 2–3 treatments were performed (SHR, 0.66; $P = 0.02$), but it was lost in the case of ≥ 4 LRTs (SHR, 0.80; $P = 0.27$). In conclusion, for MC-in patients, up to 3 LRTs are beneficial for success in intention-to-treat LT patients, with a 49% to 34% reduction in failure risk compared with untreated patients. This benefit is lost if more LRTs are required. A poor response to LRT is associated with a higher risk for HCC-dependent transplant failure.

Liver Transplantation 25 1023–1033 2019 AASLD.

Received February 15, 2019; accepted April 20, 2019.

Liver transplantation (LT) is the best curative treatment of hepatocellular carcinoma (HCC) for a cirrhotic liver.⁽¹⁾ In oncology, LT is considered a successful procedure when a

longterm posttransplant tumor-free survival is obtained. Conversely, failure is equivalent to pretransplant delisting, posttransplant tumor recurrence, or death.

Because of allograft scarcity, patients with HCC awaiting LT are most often treated using neoadjuvant locoregional therapies (LRTs) to delay tumor progression and minimize the risk of wait-list delisting.⁽²⁾ When the tumor burden meets the Milan criteria (MC) at the time of the first referral, this approach is called “bridging toward LT.” Two recent international guidelines have emphasized the importance of the

bridging strategy, for it has the potential to reduce the risk of pretransplant delisting and posttransplant recurrence. This is especially valid in cases where a partial/complete tumor response is achieved before LT.

Accordingly, LRTs have become a standard of care for treating the listed patients awaiting transplantation. However, although bridging therapies are considered a routine approach in the clinical practice, the reported quality of evidence regarding their use is poor because randomized controlled trials (RCTs) have not yet been completed.^(3,4) In this setting, RCTs are untenable due to logistical and, even more, ethical reasons. Consequently, the majority of reported studies compare posttransplant outcomes of treated and untreated patients, thereby failing to analyze the clinical course from an intention-to-treat point of view.^(5,6) Moreover, even in studies that include the wait-list period, substantial differences in tumor burden exist between initially bridged and untreated HCC patients.⁽⁷⁾

To compensate for these drawbacks, we conducted a retrospective analysis of a large European population of HCC patients meeting the MC at first referral and listed for LT. After “balancing” the cohort for possible confounders with an inverse probability of treatment weighting (IPTW), we investigated the risk factors for tumor-specific failure of LT, focusing on the role of LRT.

Patients and Methods

A large multicenter database from the European Hepatocellular Cancer and Liver Transplantation (EurHeCaLT) Study Group was initially considered for the present study. Exclusion criteria were as follows:

1. Age <18 years.
2. MC-out status at first referral.
3. Transplantation or delisting before January 1, 2001.
4. LRT other than transarterial chemoembolization (TACE), radiofrequency ablation (RFA), or percutaneous ethanol injection (PEI), ie, partial hepatectomy, transarterial radioembolization, or external radiotherapy.
5. Incidental HCC.
6. Misdiagnosed mixed hepatocellular-cholangiocellular carcinoma or cholangiocellular carcinoma.

A total of 1083 HCC patients meeting MC at first referral and enlisted for transplant during the period from January 1, 2001, to December 31, 2015, were considered for the IPTW correction. Following the correction

Abbreviations: AFP, alpha-fetoprotein; ANOVA, analysis of variance; CI, confidence interval; CR, complete response; EurHeCaLT, European Hepatocellular Cancer and Liver Transplantation; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; IPTW, inverse probability of treatment weighting; IQR, interquartile range; LRT, locoregional therapy; LT, liver transplantation; MC, Milan criteria; MELD, Model for End-Stage Liver Disease; mRECIST, modified Response Evaluation Criteria in Solid Tumors; NASH, nonalcoholic steatohepatitis; PD, progressive disease; PEI, percutaneous ethanol injection; PR, partial response; RCT, randomized controlled trial; RFA, radiofrequency ablation; SD, stable disease; SE, standard error; SHR, subhazard ratio; SMD, standardized mean difference; SPD, standardized prevalence difference; TACE, transarterial chemoembolization; WL, waiting list; WT, waiting time.

Address reprint requests to Quirino Lai, M.D., Ph.D., Starzl Abdominal Transplant Unit, Pôle de Chirurgie Expérimentale et Transplantation, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain, Avenue Hippocrates 10, 1200 Brussels, Belgium. Telephone: +32 27641412; FAX: +32 27649039; E-mail: lai.quirino@libero.it

Umberto Cillo advises Novartis and Sanofi Pharmaceuticals and receives travel grants from Johnson & Johnson, Astellas, Pfizer, Novartis, and Sanofi Pharmaceuticals. Jan Lerut has grants from Neovii Biotech and is on the speakers' bureau for Astellas. Massimo Rossi consults for Italia Pharmaceuticals.

Quirino Lai, Alessandro Vitale, and Jan Lerut were responsible for the conception, design, and analysis of the study and for writing the final report. Alessandro Vitale, Samuele Iesari, Armin Finkenstedt, Gianluca Mennini, Simona Onali, Maria Hoppe-Lotichius, Tommaso M. Manzia, Daniele Nicolini, Alfonso W. Avolio, Olga Ciccarelli, and Anna Mrzljak were involved with the collection and interpretation of data. Branislav Kocman, Salvatore Agnes, Marco Vivarelli, Giuseppe Tisone, Gerd Otto, Emmanuel Tsochatzis, Massimo Rossi, Alessandro Viveiros, Umberto Cillo, and Jan Lerut participated in data management and in the review and editing of the manuscript.

A list of European Hepatocellular Cancer and Liver Transplantation Study Group collaborators is available in the Supporting Materials.

Copyright © 2019 by the American Association for the Study of Liver Diseases.

Additional supporting information may be found in the online version of this article.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/lt.25492

for several possible confounders, a pseudo-population of 2019 patients was obtained. The study protocol received a priori approval by the Institutional Review Committee of the Université Catholique de Louvain. The study was registered at www.clinicaltrials.gov (identification number NCT03723304).

STATISTICAL ANALYSIS

Continuous variables were reported as medians and interquartile ranges (IQRs). Categorical variables were described as numbers and percentages. Comparisons between groups were made using Fisher's exact test or chi-square test for categorical variables, as appropriate. Student *t* test was used for continuous variables. Missing data relative to study covariates always involved <10% of patients. Missing data were handled using the maximum likelihood estimation method.⁽⁸⁾

To compensate for the nonrandomized design of this retrospective study, an IPTW was computed.⁽⁹⁾ The primary goal of IPTW was to achieve causal inference of an intervention (in this case, treating a patient with a LRT or not). In other terms, the IPTW analysis created a weighted sample, in which the distribution of confounding variables or prognostically important covariates was similar between treated and untreated subjects. We decided to adopt the IPTW instead of a propensity score match with the intent to avoid a reduction in the general number of the investigated population. A detailed description of the statistical strategies implemented for the IPTW construction is reported in the Supporting Materials.

Briefly, we generated a propensity score for each patient on the original population of 1083 patients. The score was created using a multivariate logistic regression model considering LRT (no versus yes) as the dependent variable. After testing several variables, we selected 8 possible clinical relevant confounders as covariates: age, sex, Model for End-Stage Liver Disease (MELD), hepatitis C virus (HCV)– and hepatitis B virus (HBV)–positive status, diameter of the largest lesion, number of nodules, and alpha-fetoprotein (AFP) level. All covariates were available at the first referral to avoid the risk of a possible immortal time bias in covariate selection. Calculating the inverse of the obtained propensity score, we “weighted” each patient and generated a pseudo-population of 2019 patients balanced for different confounders available at first referral. For example, if a patient presented an inverse propensity score number of 5, it was artificially “duplicated” for 5 times. Several tests were used to

confirm the effect of balancing. For continuous variables, we used the analysis of variance (ANOVA) *F* test and Student *t* test; for dichotomous variables, we used Fisher's exact test.

A multivariate competing risk analysis was run in the IPTW-derived population. Three different competing classes were defined:

1. Curative treatment, consisting of patients who survived after LT without recurrence.
2. HCC-dependent failure, defined as patients who exhibited tumor-related delisting or posttransplant recurrence.
3. HCC-unrelated failure, defined as the sum of patients who dropped out before LT or died after LT for causes other than HCC.

The competing risk model of HCC-dependent failure was assessed. HCC-dependent delisting was defined as any event of delisting or death on the waiting list (WL) caused by tumor progression. Immediate liver decompensation after any HCC treatment, causing death on the WL, was also considered as an HCC-dependent cause. Subhazard ratios (SHRs) and 95% CIs were reported for significant variables.⁽¹⁰⁾

Survival analyses were performed using the Kaplan-Meier method, and the log-rank test was adopted to compare the survival distributions of the examined groups.

Variables with a $P < 0.05$ were considered statistically significant. Statistical analyses were run using the SPSS, version 24.0 (SPSS Inc., Chicago, IL) and STATA, version 14.0 (StataCorp LLC, College Station, TX).

Results

COMPARISON OF TREATED VERSUS UNTREATED PATIENTS BEFORE AND AFTER IPTW

Before the IPTW, the no-LRT and LRT groups contained 182 and 901 patients ($n = 1083$). After mathematical balancing, a pseudo-population of 2019 patients was created (no LRT, $n = 1077$; LRT, $n = 942$).

Before the IPTW, several differences between the 2 groups were observed, such as higher MELD scores ($P < 0.001$) in directly transplanted patients, maximum diameter of target lesion ($P < 0.001$), and number of nodules ($P = 0.048$) in LRT patients (Table 1). A more detailed report of the differences between the 2 groups before the IPTW is shown in Supporting Table 1.

TABLE 1. Patient-, Tumor- and Transplant-Related Characteristics of the Pre-IPTW Population

Variables	No-LRT Group (n = 182)	LRT Group (n = 901)	P Value
Age at HCC diagnosis, years*	58 (51-62)	58 (52-63)	0.70
Sex, male*	153 (84.1)	741 (82.2)	0.59
On waiting list during first era (2001-2009)	87 (47.8)	318 (35.3)	0.002
Cause of cirrhosis†			
HCV*	79 (43.4)	415 (46.1)	0.57
HBV*	40 (22.0)	155 (17.2)	0.14
Alcohol	60 (33.0)	306 (34.0)	0.86
NASH	8 (4.4)	62 (6.9)	0.25
Laboratory MELD at HCC diagnosis*	12 (11-17)	12 (9-15)	<0.001
Radiological tumor features at diagnosis			
Maximum diameter of target lesion, cm*	2.0 (1.4-2.5)	2.4 (1.8-3.0)	<0.001
Number of lesions*	1 (1-2)	1 (1-2)	0.048
Radiological tumor features at LT or delisting			
Maximum diameter of target lesion, cm	2.0 (1.4-2.8)	1.7 (0.0-2.5)	<0.001
Number of lesions	1 (1-2)	1 (0-2)	0.27
AFP, ng/mL			
At tumor diagnosis*	8.3 (4.0-24.3)	9.4 (4.3-32.6)	0.32
At LT or delisting	8.7 (4.1-33.6)	8.7 (4.0-29.0)	0.48
WT duration, months	3.3 (1.1-9.1)	5.3 (2.3-10.6)	0.57
Post-LRT radiological response at LT or delisting			
CR	0	237 (26.3)	—
PR	0	258 (28.6)	—
SD	0	147 (16.3)	—
PD	0	259 (28.7)	—
Pathological tumor features‡			
Maximum diameter of target lesion, cm	2.0 (1.2-3.0)	2.0 (1.3-3.0)	0.40
Number of lesions	1 (1-2)	1 (1-3)	0.72
MC-out status	30 (17.8)	180 (23.7)	0.10
Multifocality	64 (37.9)	344 (45.4)	0.09
Poor tumor grading	21 (12.4)	105 (13.9)	0.71
Microvascular invasion	31 (18.3)	127 (16.8)	0.65
Post-LRT pathological CR	0	79 (10.4)	—

NOTE: Data are given as median (IQR) or n (%).

*Variables used for performing IPTW.

†Multiple pathologies in different patients.

‡Calculated for 927 transplanted patients in the No-LRT (n = 169) and the LRT (n = 758) groups.

After the weighting, all these variables were balanced, and several statistical tests, such as standardized differences, ANOVA F test, Fisher's exact test, and Student *t* test, confirmed the results (Tables 2 and 3). Patient-, tumor- and transplant-related characteristics of the post-IPTW pseudo-population are shown in Table 4. The only difference between the 2 post-IPTW groups was a greater median radiological dimension of the target lesion at the time of delisting/LT in the no-LRT group (2.0 versus 1.7 cm; *P* < 0.001).

The median follow-up of the investigated population from the time of first referral was 3.4 years (IQR,

1.3-7.1 years). There were 736 of 942 (78.1%) LRT patients who received TACE; 406 (43.1%) received RFA or PEI; and 200 (21.2%) patients received multiple types of treatment. For the number of treatments, 387 of 942 (41.1%) patients had 1 LRT; 346 (36.7%) had 2 or 3; and 209 (22.2%) had ≥4 treatments.

FAILURE RATES IN THE PRE- AND POST-IPTW POPULATIONS

In the original population of 1083 patients, 455 (42.0%) patients presented a failure event from the first referral

TABLE 2. Baseline Characteristics for Continuous Variables of LRT and No-LRT Patients in the Post-IPTW Population

Variable by Group	Mean	SD	SE	Variance	SMD	ANOVA F Test	Student <i>t</i> Test
Age at listing					-3.7	0.73	0.73
No LRT	56.8	8.0	0.2	64.7			
LRT	56.5	8.1	0.2	66.1			
Laboratory MELD score					-5.6	0.12	0.12
No LRT	13.5	5.0	0.2	25.4			
LRT	13.2	5.7	0.2	32.9			
Major tumor diameter					0	0.76	0.76
No LRT	2.4	1.0	0.03	1.0			
LRT	2.4	0.9	0.03	0.9			
Number of nodules					0	0.25	0.25
No LRT	1.5	0.8	0.02	0.6			
LRT	1.5	0.7	0.02	0.5			
Log ₁₀ AFP, ng/mL					0.6	0.19	0.19
No LRT	1.9	2.7	1.2	5.4			
LRT	1.9	2.6	1.1	5.2			

TABLE 3. Baseline Characteristics for Dichotomous Variables of LRT and No-LRT Patients in the Post-IPTW Population

Variable by Group	Prevalence	SPD	Fisher's Exact Test
Sex, male		-4.7	0.19
No LRT	84.6		
LRT	82.4		
HCV-positive status		-2.6	0.47
No LRT	46.6		
LRT	45.0		
HBV-positive status		1.3	0.73
No LRT	17.4		
LRT	18.0		

period to the last follow-up; the failure was HCC-related in 188 (17.4%) and unrelated to HCC in 267 (24.7%) patients. The failure caused by a delisting event was observed in 156 (14.4%) patients. HCC-dependent delisting was observed in 97 (9.0%) patients. There were 38 (3.5%) patients who died during the waiting time (WT) due to non HCC-dependent causes; the remaining 21 patients were delisted due to worsened liver function unrelated to the LRT. The failure caused by post-LT recurrence or death was reported in 299 (27.6%) patients. A total of 91 (8.4%) recurrences were reported after a median time from LT to recurrence of 22 months (IQR, 11-47 months), and 32 relapsed patients were still alive at the last follow-up. In addition, 208 (19.2%) patients died after LT due to HCC-unrelated causes (Supporting Table 2).

In the pseudo-population of 2019 post-IPTW patients, 813 (40.3%) patients presented a failure event from the first referral period to the last follow-up; the failure was HCC-related in 350 (17.3%) and unrelated to HCC in 463 (22.9%) patients. The failure caused by a delisting event was observed in 228 (11.3%) patients. HCC-dependent delisting was observed in 146 (7.2%) patients, and 45 (2.2%) patients died during the WT due to non HCC-dependent causes. The remaining 37 patients were delisted due to worsened liver function unrelated to the LRT. The failure caused by post-LT recurrence or death was reported in 585 (29.0%) patients, and 204 (10.1%) recurrences were reported after a median time from LT to recurrence of 22 months (IQR, 12-47 months). At the last follow-up, 74 relapsed patients were still alive; 381 (18.9%) patients died after LT due to HCC-unrelated causes (Supporting Table 2).

When the number of LRTs was investigated in terms of HCC-related failure rates, we performed 2 separate analyses on the pre- and post-IPTW populations. In the pre-IPTW population, patients receiving no LRT presented similar results compared with patients treated with 1 LRT (log-rank $P = 0.48$) or 2-3 treatments ($P = 0.16$). However, when the number of treatments was ≥ 4 , the failure rates grew accordingly (5-year failure rate: 30.8% versus 12.9%; $P = 0.001$; Fig. 1A).

When the post-IPTW population was investigated, we observed similar results. Patients receiving no LRT

TABLE 4. Patient-, Tumor- and Transplant-Related Characteristics of the Post-IPTW Population

Variables	No-LRT Group (n = 1077)	LRT Group (n = 942)	P Value
Age at HCC diagnosis, years*	58 (52-63)	58 (52-63)	0.73
Sex, male*	911 (84.6)	776 (82.4)	0.19
On waiting list during first era (2001-2009)	423 (39.3)	349 (37.0)	0.31
Cause of cirrhosis†			
HCV*	502 (46.6)	424 (45.0)	0.47
HBV*	187 (17.4)	170 (18.0)	0.73
Alcohol	387 (35.9)	321 (34.1)	0.40
NASH	60 (5.6)	62 (6.6)	0.35
Laboratory MELD at HCC diagnosis*	12 (11-16)	12 (9-15)	0.12
Radiological tumor features at diagnosis			
Maximum diameter of target lesion, cm*	2.2 (1.8-3.0)	2.3 (1.7-3.0)	0.75
Number of lesions*	1 (1-2)	1 (1-2)	0.25
Radiological tumor features at LT or delisting			
Maximum diameter of target lesion, cm	2.0 (1.6-3.0)	1.7 (0.0-2.4)	<0.001
Number of lesions	1 (1-2)	1 (0-2)	0.80
AFP, ng/mL			
At tumor diagnosis*	9.2 (4.0-26.8)	9.2 (4.2-32.7)	0.81
At LT or delisting	10.0 (4.0-39.0)	8.7 (4.0-29.2)	0.33
WT duration, months	2.9 (1.0-8.7)	5.2 (2.2-10.6)	0.57
Post-LRT radiological response at LT or delisting			
CR	0	253 (26.9)	—
PR	0	263 (27.9)	—
SD	0	147 (15.6)	—
PD	0	275 (29.2)	—
Pathological tumor features‡			
Maximum diameter of target lesion, cm	2.3 (1.4-3.0)	2.0 (1.3-3.0)	0.11
Number of lesions	1 (1-2)	1 (1-3)	0.48
MC-out status	213 (21.2)	185 (23.6)	0.23
Multifocality	435 (43.2)	357 (45.5)	0.36
Poor tumor grading	141 (14.0)	107 (13.6)	0.84
Microvascular invasion	197 (19.6)	134 (17.1)	0.18
Post-LRT pathological CR	0	81 (10.3)	—

NOTE: Data are given as median (IQR) or n (%).

*Variables used for performing IPTW.

†Multiple pathologies in different patients.

‡Calculated for 1791 transplanted patients in the No-LRT (n = 1007) and the LRT (n = 784) groups.

presented similar results compared with patients treated with 1 LRT (log-rank $P = 0.77$) or 2-3 treatments ($P = 0.10$). However, when the number of treatments was ≥ 4 , the failure rates grew accordingly (5-year failure rate: 31.6% versus 15.9%; $P < 0.001$; Fig. 1B).

RISK FACTORS OF HCC-DEPENDENT FAILURE

The risk factors for the competing event of HCC-dependent failure in the pre- and post-IPTW

populations are shown in Table 5. In the pre-IPTW population, progressive tumor disease at last radiological assessment was the most important independent risk factor for HCC-dependent failure (SHR, 5.70; $P < 0.001$), followed by the AFP level (SHR, 1.53; $P < 0.001$) at first referral. Other significant independent tumor-related risk factors were the diameter of the major lesion, MELD, patient age, and period of WL registration. Interestingly, no statistically significant effect was reported concerning the number of LRTs performed. Only receiving 1 LRT

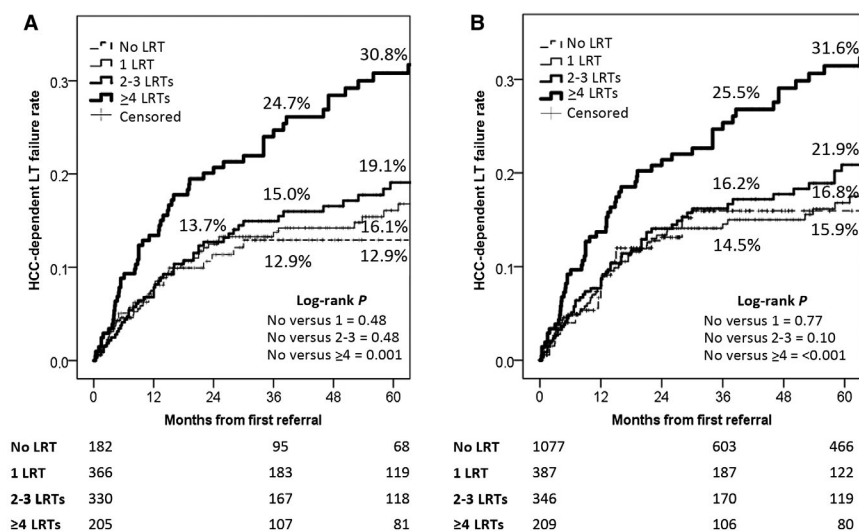


FIG. 1. (A). HCC-dependent failure rates in the pre-IPTW population stratified by number of LRTs. (B) HCC-dependent failure rates in the post-IPTW population stratified by number of LRTs.

TABLE 5. Competing Risk Model for the Risk of HCC-Dependent LT Strategy Failure Before and After IPTW

Variables	β	SE	SHR	95%CI	P Value
Before IPTW					
On waiting list during first era (2001-2009)	0.36	0.17	1.43	1.02-1.99	0.04
Sex, male	0.24	0.21	1.26	0.84-1.91	0.26
Age at first referral	0.03	0.01	1.03	1.01-1.05	0.01
MELD at first referral	0.03	0.01	1.03	1.00-1.06	0.02
Diameter of largest tumor at first referral	0.17	0.08	1.19	1.02-1.38	0.03
LRT					
No LRT	Reference	—	1.00	—	—
1 LRT	-0.54	0.28	0.58	0.34-1.00	0.05
2-3 LRTs	-0.32	0.27	0.73	0.43-1.23	0.24
≥4 LRTs	-0.09	0.28	0.92	0.53-1.60	0.76
$\log_{10}$ AFP at first referral	0.42	0.10	1.53	1.26-1.86	<0.001
mRECIST PD	1.74	0.17	5.70	4.11-7.90	<0.001
After IPTW					
On waiting list during first era (2001-2009)	0.42	0.12	1.52	1.20-1.91	<0.001
Sex, male	0.55	0.18	1.73	1.23-2.44	0.002
Age at first referral	0.42	0.07	1.53	1.34-1.74	<0.001
MELD at first referral	0.03	0.01	1.03	1.01-1.06	0.002
Diameter of largest tumor at first referral	0.27	0.05	1.31	1.18-1.45	<0.001
LRT					
No LRT	Reference	—	1.00	—	—
1 LRT	-0.67	0.18	0.51	0.36-0.74	<0.001
2-3 LRTs	-0.42	0.18	0.66	0.47-0.93	0.02
≥4 LRTs	-0.22	0.20	0.80	0.55-1.17	0.27
$\log_{10}$ AFP at first referral	0.48	0.07	1.62	1.41-1.87	<0.001
mRECIST PD	1.73	0.16	5.62	4.10-7.69	<0.001

merged statistical significance, presenting a trend for protection with respect to direct LT (SHR, 0.58; $P = 0.05$).

In the post-IPTW population, progressive tumor disease at last radiological assessment was the most important independent risk factor for HCC-dependent failure (SHR, 5.62; $P < 0.001$). The diameter of the major lesion (SHR, 1.31; $P < 0.001$) and AFP level (SHR, 1.62; $P < 0.001$) at first referral were tumor-related risk factors. MELD, patient age, male sex, and period of WL registration also directly influenced the risk of HCC-dependent failure. Patients who received only 1 LRT had the best protective effect against failure compared with untreated patients (SHR, 0.51; $P < 0.001$). This beneficial effect was apparent as long as 2-3 treatments were done (SHR, 0.66; $P = 0.02$), but it was lost in case of further treatments (SHR, 0.80; $P = 0.27$). Supporting Table 4 shows the different effects of the investigated risk factors in the 2 separate components of HCC-related failure, namely, delisting and recurrence.

Discussion

In oncology, establishing the superiority of one therapeutic strategy over another one requires RCTs, which aim to identify proportions of therapeutic failure (ie, progressive disease [PD], recurrence, or death) between the 2 approaches.⁽¹¹⁾ However, as is usually the case, basic oncological principles are overlooked in the field of transplant oncology. Given the shortcomings of statistical methodology, 3 different reasons might explain the lack of clarity about the LRT effect as a neoadjuvant treatment in LT.

First, an RCT that compares patients receiving upfront transplants with patients receiving LRT first as a bridge to LT is difficult to support from an ethical point of view. Thus, we developed a sophisticated statistical approach with the intent to balance a historical population of no-LRT and LRT patients on the basis of information available at the first referral for LT. The IPTW strongly affected sample size, by “artificially” increasing the number of no-LRT patients. Nevertheless, such methodology was the only key to offset the important, otherwise unresolvable, initial selection bias. The balancing effect should be noted observing how the results of the competing risk analyses changed in the pre- and post-IPTW population: This phenomenon was the consequence of the limitation

of the initial biases presented in the original population. The IPTW is prone to conceptual drawbacks, but this methodology represents the most rigorous way to re-equilibrate the sample to test.⁽¹²⁾ Consistently, all the tests used to check the successful balancing of the 2 study groups confirmed the validity of our method.

Second, studies comparing no-LRT and LRT patients focus only on posttransplant data, thereby failing to obtain intention-to-treat results. It is only recently that the importance of intention-to-treat analyses in the setting of LT has been recognized.⁽¹³⁻¹⁶⁾ For the first time, the present study has investigated the intention-to-treat effect of LRT against upfront LT in MC-in HCC patients.

Third, the overlapping effects of several risk factors might lead to inaccurate results because of the absence of a competing risk analysis. A competing risk is an event that either hides the observation of the event in the study (ie, HCC-related outcomes) or alters the chance that this event occurs. Recently, the statistical analysis of competing risks has also been introduced in the setting of HCC and LT.⁽¹⁷⁾ Indeed, the competing risk analysis brings about the definition of real-world probabilities of a specific event by breaking down specific causes.

In our study, 2 risks that compete with curative treatment were considered: failure caused by tumor-related reasons (ie, pretransplant delisting caused by disease progression and posttransplant recurrence) and failure caused by nontumor-related events. The conceptual evaluation of pretransplant and posttransplant adverse events through the same “failure approach” represents a novelty in the LT set. In this analysis, disease progression represented the worst independent risk factor with a 5.62-fold increased risk for HCC-dependent failure. This observation is in line with many studies showing the detrimental role of poor radiological response on delisting, intention-to-treat death, transplant benefit, posttransplant tumor recurrence, and posttransplant death.^(5,7,13,14,18,19) The diameter of the largest lesion, AFP levels at first referral, and MELD scores were also risk factors for HCC-dependent failure by previous reports.^(2,7,13-23)

This study revealed that up to 3 pretransplant LRTs were protective compared with no LRT. One LRT reduced the risk of intention-to-treat failure by 49%, and 2 to 3 treatments decreased this risk by 34%. These findings are in line with the recent international guidelines, which suggest that bridging LRTs are appropriate in a LT project, despite the heterogeneity of the reported data.^(3,4)

Interestingly, when we investigated the risk factors for HCC-related failure in the pre-IPTW population, LRT number failed to be statistically significant. This result suggests that the investigation of the LRT effect on delisting and posttransplant recurrence should be markedly influenced by the initial heterogeneity of the investigated population. A recent meta-analysis, focusing on LRT and LT, specifically pointed out the heterogeneity biases among different studies, which are caused by variable demographics (ie, the great variability of WT), types of LRT, HCC stages (T1 versus T2), and treatment dynamics (frequency and interval between treatments).⁽²⁴⁾ Despite these limitations, that meta-analysis partially hinted at our results: LRT proved beneficial in terms of global delisting rates (relative risk, 0.19; 95% CI, 0.15-0.24) and HCC-dependent delisting (relative risk, 0.11; 95% CI, 0.07-0.17).⁽²⁴⁾ When studies comparing treated and untreated patients were tested, the relative risk seemed protective (0.32; 95% CI, 0.06-1.85). Nonetheless, the effect was not statistically significant, probably because of biases, imprecision, and inconsistency in the included studies.⁽²⁴⁾ The beneficial effect of LRT on the risk of delisting has also been reported in the Western and Eastern literature.^(7,14,22,23)

The positive effect of upfront LRT was described in the recent multicenter US experience comprising 3601 patients reported by Agopian et al.⁽⁵⁾ One LRT was protective for the risk of recurrence compared with a direct transplant (HR, 0.86); conversely, an increased number of treatments raised the risk (≥ 4 LRT: HR, 2.5; $P < 0.001$).

Our results concerning LRT in MC-in HCC patients might explain the discrepancy within previous reports. The different number of bridge treatments determines different pretransplant and posttransplant outcomes. The decision to perform a direct transplant shifts the risk of pretransplant delisting into the risk of posttransplant recurrence by eliminating LRT as a selection criterion. This phenomenon is also shown in Supporting Table 4, in which a higher percentage of no-LRT patients experienced recurrence with respect to treated patients, whereas LRT patients presented more cases of PD and longer WTs (namely, selection by time and LRT). This argument has also been clearly shown in “fast-track” living donor LT and in studies about WT as a possible selector for the risk of posttransplant recurrence.⁽²⁵⁻²⁷⁾ In all of these studies, the patients presenting shorter WTs also received a lower number of LRTs.

Our study shows that performing a pretransplant LRT strategy gives a beneficial effect on post-LT results as long as it does not exceed 3 treatments. In other terms, when the number of required LRTs for stabilizing the tumor is ≤ 3 , the positive effect (namely, reducing the posttransplant recurrence) is statistically and significantly superior with respect to the negative one (namely, increasing the delisting rates). When ≥ 4 bridging treatments are necessary with the intent to stabilize the tumor before LT, this positive phenomenon disappears. This negative course is a clear demonstration of a tumor selection by treatment: the higher the number of LRTs required, the higher the aggressiveness of the tumor and the worse the overall results.

Very low 5-year HCC-dependent failure rates can be achieved in patients who initially received 1 LRT and did not exhibit disease progression after the treatment. Similar data were also observed in a large retrospective US analysis performed on 2794 LT patients, in which a lower posttransplant recurrence rate was reported in patients undergoing LRT, whereas AFP and tumor burden were independent risk factors for recurrence.⁽²⁸⁾ In light of these results, we can postulate that tumor characteristics prevail over the treatment in influencing the ultimate therapeutic results. Still, using LRT as a selection tool strongly discriminates between patients in terms of posttransplant clinical course.

In an era in which great pressure exists on health care quality improvement and cost reduction, our study suggests the opportunity to frame a shift in standard practice toward LRT in MC-in HCC patients. In our study, we observed a risk reduction from 34% to 49% in patients receiving ≤ 3 LRTs, possibly resulting in a substantial, cost-effective benefit. However, large prospective cost-effective analyses are needed to confirm this effect. Response to LRT seems to be a rather rudimentary but valuable predictive tool, being able to unveil tumor biology and, as a test of time, to select patients for LT. However, the decision to incorporate LRT as a standard approach in MC-in patients should also be implemented in light of different local philosophies. As an example, it could be argued that LRT could be offered as a standard approach in patients with a predicted WT ≥ 6 months. On the opposite side, studies clarifying the benefit of LRT in centers with shorter median times or with living donor programs require further evaluation.

The specific role of the LRT method used in our series has been only marginally explored in the study. In Supporting Table 5, we reported the different risks

of delisting, recurrence, and overall HCC-dependent failure according to the use of an embolic versus ablative versus combined approach. Preliminary evidence exists on the necessity of treating tumors with a combinatory approach during the WT, and how it should be connected with an increased risk of overall tumor-related failure. However, we should consider these preliminary results with caution. Further analyses that focus on this aspect are required.

Our study has some limitations. The retrospective nature of the study impaired our possibility to specify in detail the reasons justifying a repetitive treatment approach (ie, LRT refractoriness of the target lesion versus initial multimodal approach versus treatment of new tumors).

Moreover, the study, which covers a long period, implies an evolution in the technical aspects of LRT. Hence, the study period was limited to the last 2 decades, and the period was introduced as a variable in the multivariate model to correct the results also for this possible confounder. The analysis showed that patients listed during the first era (2001–2009) fared worse, possibly because therapeutic approaches improved during the second era. One can postulate that excluding the cases of hepatic resection, transarterial radioembolization, or external radiotherapy should reduce the impact of our intention-to-treat analysis. On the other hand, we think that considering salvage LT after resection or very uncommon strategies, such as radioembolization or radiotherapy, should represent a bias, mainly due to their neglectable number in the present study.

Another possible limitation concerns intercenter differences in the length of WT and dynamics of LRT. Although these discrepancies are difficult to resolve, the centers belonging to the EurHeCaLT Study Group adopt similar approaches and policies in the HCC management before LT. Moreover, the difference across centers can be statistically beneficial because it enriches the variability within patient cohorts and brings about more solid statistical results.

Lastly, the use of the IPTW can be criticized for the artificial increase in the sample size. However, this sophisticated statistical approach is the only way forward in removing the initial selection bias. We are unable to assert that any residual confounding may occur in the study because of an imperfect measure of some confounder initially used for the construction of the IPTW model. Unfortunately, when this phenomenon is observed, an adjustment done using this imperfect measure does not completely remove the effect of

the confounding variable. We should honestly underline that this latter limit, which derives from the great initial difference among LRT and no-LRT patients, makes it very difficult to construct a balanced model and that it is probably impossible to eliminate in this specific setting.

In conclusion, LRT for MC-in HCC patients is valuable when considering an intention-to-treat LT approach. When comparing treated and untreated patients, 1 single and 2–3 LRTs confer a 49% and 34% reduction, respectively, in the risk of HCC-dependent transplant failure. This beneficial effect disappears with ≥ 4 LRTs because of more aggressive tumor biology. Patients who show a poor response to LRT have a predictably greater risk for pretransplant tumor-related delisting or posttransplant recurrence.

Acknowledgments: The authors thank David Joseph Tannert and Marta Fiorenza for their linguistic revision.

REFERENCES

- 1) Lai Q, Lerut JP. Hepatocellular cancer: how to expand safely inclusion criteria for liver transplantation. *Curr Opin Organ Transplant* 2014;19:229–234.
- 2) Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med* 1996;334:693–699.
- 3) European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018;69:182–236.
- 4) Heimbach JK, Kulik LM, Finn RS, Sirlin CB, Abecassis MM, Roberts LR, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology* 2018;67:358–380.
- 5) 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.
- 6) Terzi E, Ray Kim W, Sanchez W, Charlton MR, Schmeltzer P, Gores GJ, et al. Impact of multiple transarterial chemoembolization treatments on hepatocellular carcinoma for patients awaiting liver transplantation. *Liver Transpl* 2015;21:248–257.
- 7) Mehta N, Dodge JL, Goel A, Roberts JP, Hirose R, Yao FY. Identification of liver transplant candidates with hepatocellular carcinoma and a very low dropout risk: implications for the current organ allocation policy. *Liver Transpl* 2013;19:1343–1353.
- 8) Kang S, Little RJ, Kaciroti N. Missing not at random models for masked clinical trials with dropouts. *Clin Trials* 2015;12:139–148.
- 9) Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 2015;34:3661–3679.

- 10) Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999;94: 496-509.
- 11) Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, et al.; for RESORCE Investigators. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 2017;389:56-66.
- 12) Li F, Thomas LE, Li F. Addressing extreme propensity scores via the overlap weights. *Am J Epidemiol* 2019;188:250-257.
- 13) Lai Q, Nicolini D, Inostroza Nunez M, Iesari S, Goffette P, Agostini A, et al. A novel prognostic index in patients with hepatocellular cancer waiting for liver transplantation: time-radiological-response-alpha-fetoprotein-inflammation (TRAIN) score. *Ann Surg* 2016;264:787-796.
- 14) Lai Q, Vitale A, Iesari S, Finkenstedt A, Mennini G, Spoleitini G, et al.; for European Hepatocellular Cancer Liver Transplant Study Group. Intention-to-treat survival benefit of liver transplantation in patients with hepatocellular cancer. *Hepatology* 2017;66:1910-1919.
- 15) Lee DD, Samoylova M, Mehta N, Musto KR, Roberts JP, Yao FY, Harnois DM. The mRECIST classification provides insight into tumor biology for patients with hepatocellular carcinoma awaiting liver transplantation. *Liver Transpl* 2019;25:228-241.
- 16) Firl DJ, Kimura S, McVey J, Hashimoto K, Yeh H, Miller CM, et al. Reframing the approach to patients with hepatocellular carcinoma: longitudinal assessment with hazard associated with liver transplantation for HCC (HALTHCC) improves ablate and wait strategy. *Hepatology* 2018;68:1448-1458.
- 17) 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.
- 18) Mehta N, Sarkar M, Dodge JL, Fidelman N, Roberts JP, Yao FY. Intention to treat outcome of T1 hepatocellular carcinoma with the "wait and not ablate" approach until meeting T2 criteria for liver transplant listing. *Liver Transpl* 2016;22:178-187.
- 19) Lai Q, Avolio AW, Graziadei I, Otto G, Rossi M, Tisone G, et al.; for European Hepatocellular Cancer Liver Transplant Study Group. Alpha-fetoprotein and modified response evaluation criteria in solid tumors progression after locoregional therapy as predictors of hepatocellular cancer recurrence and death after transplantation. *Liver Transpl* 2013;19:1108-1118.
- 20) Duvoux C, Roudot-Thoraval F, Decaens T, Pessione F, Badran H, Piardi T, et al.; for Liver Transplantation French Study Group. Liver transplantation for hepatocellular carcinoma: a model including α -fetoprotein improves the performance of Milan criteria. *Gastroenterology* 2012;143:986-994.
- 21) 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.
- 22) Huo TI, Huang YH, Su CW, Lin HC, Chiang JH, Chiou YY, et al. Validation of the HCC-MELD for dropout probability in patients with small hepatocellular carcinoma undergoing locoregional therapy. *Clin Transplant* 2008;22:469-475.
- 23) Park SJ, Freise CE, Hirose R, Kerlan RK, Yao FY, Roberts JP, Vagefi PA. Risk factors for liver transplant waitlist dropout in patients with hepatocellular carcinoma. *Clin Transplant* 2012;26:E359-E364.
- 24) Kulik L, Heimbach JK, Zaiem F, Almasri J, Prokop LJ, Wang Z, et al. Therapies for patients with hepatocellular carcinoma awaiting liver transplantation: a systematic review and meta-analysis. *Hepatology* 2018;67:381-400.
- 25) Halazun KJ, Patzer RE, Rana AA, Verna EC, Griesemer AD, Parsons RF, et al. Standing the test of time: outcomes of a decade of prioritizing patients with hepatocellular carcinoma, results of the UNOS natural geographic experiment. *Hepatology* 2014;60:1957-1962.
- 26) Samoylova ML, Dodge JL, Yao FY, Roberts JP. Time to transplantation as a predictor of hepatocellular carcinoma recurrence after liver transplantation. *Liver Transpl* 2014;20:937-944.
- 27) Lai Q, Avolio AW, Lerut J, Singh G, Chan SC, Berloco PB, et al. Recurrence of hepatocellular cancer after liver transplantation: the role of primary resection and salvage transplantation in East and West. *J Hepatol* 2012;57:974-979.
- 28) Oligane HC, Xing M, Kim HS. Effect of bridging local-regional therapy on recurrence of hepatocellular carcinoma and survival after orthotopic liver transplantation. *Radiology* 2017;282:869-879.