DR. QUIRINO LAI (Orcid ID : 0000-0003-1487-3235)

DR. TOMMASO M MANZIA (Orcid ID : 0000-0002-4636-3478)

DR. EMMANUEL A TSOCHATZIS (Orcid ID : 0000-0001-5069-2461)

Article type : Original Articles

THE INTENTION-TO-TREAT EFFECT OF BRIDGING TREATMENTS IN THE SETTING OF MILAN CRITERIA-IN PATIENTS WAITING FOR LIVER TRANSPLANTATION

Quirino LAI, MD PhD,[1,2] Alessandro VITALE, MD PhD,[3] Samuele IESARI, MD,[1,4] Armin FINKENSTEDT, MD,[5] Gianluca MENNINI, MD PhD,[2] Simona ONALI, MD,[6] Maria HOPPE-LOTICHIUS, MD,[7] Tommaso M. MANZIA, MD PhD,[8] Daniele NICOLINI, MD PhD,[9] Alfonso W. AVOLIO, MD,[10] Anna MRZLJAK, MD,[11] Branislav KOCMAN, MD,[11] Salvatore AGNES, MD,[10] Marco VIVARELLI, MD PhD,[9] Giuseppe TISONE, MD,[8] Gerd OTTO, MD,[7] Emmanuel TSOCHATZIS, MD PhD,[6] Massimo ROSSI, MD,[2] Andre VIVEIROS, MD,[5] Olga CICCARELLI, MD,[1] Umberto CILLO, MD,[3] Jan LERUT, MD PhD;[1] on behalf of the European Hepatocellular Cancer Liver Transplant Study Group

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

Quirino Lai: lai.quirino@libero.it
Alessandro Vitale: alessandro.vitale@unipd.it
Samuele Iesari: samuele.iesari@uclouvain.be
Armin Finkenstedt: armin.finkenstedt@tirol-kliniken.at
Gianluca Mennini: gianluca.mennini@uniroma1.it
Simona Onali: s.onali@ucl.ac.uk
Maria Hoppe-Lotichius: maria.hoppe-lotichius@unimedizin-mainz.de
Tommaso M Manzia: tomanzia@libero.it
Daniele Nicolini: nicolini_daniele@yahoo.it
Alfonso W Avolio: alfonso.avolio@policlinicogemelli.it
Anna Mrzljak: anna.mrzljak@mef.hr
Branislav Kocman: branislav.kocman@kb-merkur.hr
Salvatore Agnes: salvatoreagnes@policlinicogemelli.it
Marco Vivarelli: vivarelli63@libero.it
Giuseppe Tisone: tisone@med.uniroma2.it
Gerd Otto: gerd.otto@unimedizin-mainz.de
Emmanuel Tsochatzis: e.tsochatzis@ucl.ac.uk
Massimo Rossi: massimo.rossi@uniroma1.it
Andre Viveiros: andre.viveiros@i-med.ac.at
Olga Ciccarelli: olga.ciccarelli@uclouvain.be
Umberto Cillo: cillo@unipd.it
Jan Lerut: jan.lerut@uclouvain.be

Keywords
Liver transplantation; hepatocellular cancer; locoregional therapy; alpha-fetoprotein; mRECIST.

Address reprint requests to:
Quirino LAI, M.D., Ph.D.
Jan LERUT, M.D., Ph.D.,
Starzl Abdominal Transplant Unit, University Hospitals Saint Luc,
Université catholique de Louvain (UCL),
Avenue Hippocrates 10, 1200 Brussels, Belgium;
Phone: +32 27641412; fax: +32 27649039.
E-mails: lai.quirino@libero.it
Jan.lerut@uclouvain.be

Conflict of interest
The authors have no conflicts of interest to declare concerning the present study.

Funds
No funds were required nor obtained for the present study.

Contributors

QL, AV, and JL were responsible for the conception, design, and analysis of the study and for writing the final report; AV, SI, AF, GM, SO, MHL, TMM, DN, AWA, OC, and AM were involved with the collection and interpretation of data; BK, SA, MV, GT, GO, ET, MR, AV, UC, and JL participated in data management, review and editing of the manuscript.

Acknowledgments

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

Abbreviations used in the text:

AFP	alpha-fetoprotein
ANOVA	analysis of variance
CI	confidence intervals
EurHeCaLT	European Hepatocellular Cancer and Liver Transplantation
HCC	hepatocellular cancer
IPTW	inverse probability of treatment weighting
IQR	interquartile ranges
LRT	loco-regional treatment
LT	liver transplantation
MC	Milan Criteria
MELD	model for end-stage liver disease
mRECIST	modified Response Evaluation Criteria in Solid Tumors
PEI	percutaneous ethanol injection
RCT	randomized controlled trial
RFA	radio-frequency ablation
SHR	sub-hazard ratios
TACE	trans-arterial chemo-embolization
WL	waiting list

ABSTRACT

In patients with hepatocellular cancer (HCC) meeting the Milan Criteria (MC), the benefit of loco-regional treatments (LRT) 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 pre-transplant tumor-related de-listing or post-transplant recurrence. The study was registered at http://www.ClinicalTrials.gov (ID:NCT03723304). In order to offset the initial limitations of the investigated population, an inverse probability of treatment weighting (IPTW) analysis was used: 1,083 MC-IN cases

(no-LRT=182; LRT=901) were balanced using eight variables: age, gender, MELD value, HCV status, HBV 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 2,019 patients listed for LT was analyzed, comparing two homogeneous groups of untreated (N=1,077) and LRT-treated (N=942) patients. Tumor progression after LRT was the most important independent risk factor for HCC-dependent failure (sub-hazard ratio=5.62; p<0.001). Other independent risk factors were major tumor diameter, AFP, MELD, patient age, male gender and period of waiting-list registration. One single LRT was protective compared with no treatment (SHR=0.51; p<0.001). The positive effect was still observed when two to three treatments were performed (SHR=0.66; p=0.02), but it was lost in case of four or more LRT (SHR=0.80; p=0.27). **CONCLUSION:** In MC-IN patients, up to three LRT are beneficial about intention-to-treat LT success, with a 49 to 34% reduction in failure risk compared to untreated patients. This benefit is lost if more LRT are required. Poor response to LRT is associated with a higher risk for HCC-dependent transplant failure.

INTRODUCTION

Liver transplantation (LT) is the best curative treatment of hepatocellular cancer (HCC) developed on a cirrhotic liver (1). In oncology, LT is considered a successful procedure when a long-term post-transplant tumor-free survival is obtained. Conversely, failure is equivalent to pre-transplant de-listing, post-transplant tumor recurrence or death.

Due to allograft scarcity, patients with HCC awaiting LT are most often treated using neo-adjuvant loco-regional treatments (LRT) to delay tumor progression and minimize the risk of wait-list de-listing. (2) When the tumor burden meets the Milan Criteria (MC) at the time of the first referral, this approach is called “bridging towards LT”. Two recent international

guidelines have emphasized the importance of the bridging strategy, for it has the potential to reduce the risk of pre-transplant de-listing and post-transplant recurrence. This is especially valid in cases where a partial/complete tumor response is achieved before LT.

Accordingly, LRT 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, since 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 compares post-transplant 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 waiting-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 liver transplantation, focusing on the role of LRT.

MATERIALS AND METHODS

A large multicentre database coming from the European Hepatocellular Cancer and Liver Transplantation (EurHeCaLT) Study Group was initially considered for the present study. Exclusion criteria were: a) age <18 years; b) MC-OUT status at first referral; c) transplantation or delisting before January 1, 2001; d) LRT other than trans-arterial chemo-

embolization (TACE), radio-frequency ablation (RFA) or percutaneous ethanol injection (PEI) (i.e. partial hepatectomy, trans-arterial radio-embolization or external radiotherapy); e) incidental HCC; and f) misdiagnosed mixed hepatocellular-cholangiocellular carcinoma or cholangiocellular carcinoma.

A total of 1,083 HCC patients meeting MC at first referral and enlisted for transplant during the period January 1, 2001 – December 31, 2015, were considered for the IPTW correction. Following the correction for several possible confounders, a pseudo-population of 2,019 patients was obtained. The study was registered at http://www.ClinicalTrials.gov (ID: NCT03723304).

Statistical analysis

Continuous variables were reported as medians and interquartile ranges (IQR). 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's t-test was used for continuous variables. Missing data relative to study covariates always involved less than 10% of patients. Missing data were handled using the maximum likelihood estimation method. (8)

To compensate for the non-randomized 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 **Additive Data** section. Briefly, we generated a propensity score for each patient on the original population of 1,083 patients. The score was created using a multivariate logistic regression model considering "loco-regional treatment" (LRT) (no vs. yes) as the dependent variable. After testing several variables, we selected eight possible clinical relevant confounders as co-variates: age, gender, model for the end-stage liver disease (MELD), hepatitis C and B 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 2,019 patients balanced for different confounders available at first referral. As an example, if a patient presented an inverse propensity score number of five, it was artificially "duplicated" for five 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's t-test; for dichotomous variables, we used Fisher's exact test.

A multivariable competing-risk analysis was run in the IPTW-derived population. Three different competing classes were defined: a) curative treatment, consisting of patients who survived after LT without recurrence; b) HCC-dependent failure, defined as patients who exhibited tumor-related de-listing or post-transplant recurrence; c) 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 de-listing was defined as any event of delisting or death on the waiting list caused by tumor progression. Immediate liver decompensation after any HCC treatment, causing

death on the waiting list, was also considered as an HCC-dependent cause. Sub-hazard ratios (SHRs) and 95%CI 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 statistical package version 24.0 (SPSS Inc., Chicago, IL, USA) and STATA statistical package version 14.0 (StataCorp LLC, College Station, TX, USA).

RESULTS

Comparison of treated vs. untreated patients before and after the IPTW

Before the IPTW, the no-LRT and LRT groups contained 182 and 901 cases (total=1,083). After mathematical balancing, a pseudo-population of 2,019 patients was created (no-LRT=1,077; LRT=942).

Before the IPTW, several differences between the two groups were observed, such as higher MELD scores ($p<0.001$) in directly transplanted patients and initially greater tumor burden ($p<0.001$) and a major number of nodules ($p=0.048$) in LRT patients (**Table 1**). A more detailed report of the differences between the two groups before the IPTW is shown in **Additive Table 1**.

After the weighting, all these variables were "balanced", and several statistical tests, such as standardized differences, ANOVA, Fisher's exact test, and Student's t-test, confirmed the results (**Table 2**). Patient-, tumor- and transplant-related characteristics of the post-IPTW pseudo-population are shown in **Table 3**. The only difference between the two post-IPTW

groups was a greater median radiological dimension of the target lesion at the time of delisting/LT in the non-LRT group (2.0 vs. 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). Seven hundred thirty-six (78.1%) of 942 LRT patients received TACE, 406 (43.1%) RFA or PEI. Two hundred (21.2%) patients received multiple types of treatment. Three hundred eighty-seven (41.1%) of 942 patients had one LRT, 346 (36.7%) two or three, and 209 (22.2%) four or more treatments.

Failure rates in the pre- and post-IPTW populations

In the original population of 1,083 cases, 455 (42.0%) patients presented a failure event from the first referral period to the last follow-up; the failure was HCC-related in 188 (17.4%) and unrelated to HCC in 267 (24.7%) cases. The failure caused by a de-listing event was observed in 156 (14.4%) patients. HCC-dependent de-listing was observed in 97 (9.0%) patients. Thirty-eight (3.5%) patients died during the waiting time 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. Ninety-one (8.4%) recurrences were reported after a median time from LT to recurrence of 22 months (IQR=11-47). Thirty-two relapsed patients were still alive at the last follow-up. Two hundred and eight (19.2%) patients died after LT due to HCC-unrelated causes (**Additive Table 2**).

In the pseudo-population of 2,019 post-IPTW cases, 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%) cases. The failure caused by a de-listing event was observed in 228 (11.3%) patients. HCC-dependent de-listing was observed in 146 (7.2%) patients. Forty-five (2.2%) patients died during the waiting time 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. Two hundred and four (10.1%) recurrences were reported after a median time from LT to recurrence of 22 months (IQR=12-47). Seventy-four relapsed patients were still alive at the last follow-up. Three hundred eighty-one (18.9%) patients died after LT due to HCC-unrelated causes (**Additive Table 2**).

When the number of LRT was investigated in terms of HCC-related failure rates, we performed two separate analyses on the pre- and post-IPTW population. In the pre-IPTW population, patients receiving no LRT presented similar results compared to subjects treated with one LRT (log-rank p=0.48), or two-to-three treatments (p=0.16). However, when the number of treatments was ≥4, the failure rates grew accordingly (five-year failure rate: 30.8% vs. 12.9%; p=0.001) (**Figure 1A**).

When the post-IPTW population was investigated, we observed similar results. Patients receiving no LRT presented similar results compared to subjects treated with one LRT (log-rank p=0.8), or two-to-three treatments (p=0.1). However, when the number of treatments was ≥4, the failure rates grew accordingly (five-year failure rate: 31.6% vs. 15.9%; p<0.001) (**Figure 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 4**. 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 LRT performed. Only receiving one LRT merged statistical significance, presenting a trend for protection respect to direct LT (SHR=0.58; p=0.050).

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 gender and period of WL registration also directly influenced the risk of HCC-dependent failure. Patients who received only one LRT had the best protective effect against failure compared to untreated cases (SHR=0.51; p<0.001). This beneficial effect was apparent as long as two to three treatments were done (SHR=0.66; p=0.02), but it was lost in case of further treatments (SHR=0.80; p=0.27). **Additive Table 4** has also been reported showing the different effects of the investigated risk factors in the two separate components of HCC-related failure, namely de-listing 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 (i.e., progressive disease, recurrence or death) between the two 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, three different reasons might explain the lack of clarity about the LRT effect as a neo-adjuvant treatment in liver transplantation.

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 cases, based on information available at the first referral for LT. The IPTW strongly affected sample size, by "artificially" increasing the number of no-LRT cases. 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 two study groups confirmed the validity of our method.

Secondly, studies comparing no-LRT and LRT patients focus only on post-transplant data, thereby failing to obtain intention-to-treat results. Only recently, 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.

Thirdly, the overlapping effects of several risk factors might lead to inaccurate results owing to the absence of competing-risk analysis. A competing risk is an event that either hides the observation of the event in the study (i.e., 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, two risks that compete with curative treatment were considered: failure caused by tumor-related reasons (i.e., pre-transplant de-listing caused by disease progression and post-transplant recurrence) and failure caused by non-tumor related events. The conceptual evaluation of pre- and post-transplant 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, post-transplant tumor recurrence, and post-transplant death. (5,7,13,14,18,19) The diameter of the largest lesion, AFP levels at first referral, and MELD were also risk factors for HCC-dependent failure, by previous reports (2,7,13-23)

This study revealed that up to three pre-transplant LRT were protective compared to no LRT. One LRT reduced the risk of intention-to-treat failure by 49%, and two to three treatments decreased this risk by 34%. These findings are in line with the recent international guidelines, which suggest that bridging LRT 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 de-listing and post-transplant recurrence should be markedly influenced by the initial heterogeneity of the investigated population. A recent meta-analysis, focusing on LRT and LT, has specifically pointed out the heterogeneity biases among different studies, caused by variable demographics (i.e., the great variability of waiting time), types of LRT, HCC stages (T1 vs. T2), and treatment dynamics (frequency and interval between treatments). (24) Despite these limitations, that meta-analysis has partially hinted our results: LRT proved beneficial in terms of global de-listing 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 cases 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 has been described in the recent multicentre US experience comprising 3,601 patients reported by Agopian et al. One LRT was protective for the risk of recurrence compared to direct transplant (HR=0.86); conversely, an increased number of treatments raised the risk (≥4 LRT: HR=2.5; $p<0.001$). (5)

Our results concerning LRT in MC-IN HCC patients might explain the discrepancy within previous reports. The different number of bridge treatments determines different pre- and post-transplant outcomes. The decision to perform a direct transplant shifts the risk of pre-transplant de-listing into the risk of post-transplant recurrence, by eliminating LRT as a selection criterion. This phenomenon is also shown in **Additive Table 4**, in which a higher percentage of no-LRT patients experienced recurrence respect to treated subjects, while LRT patients presented more cases of progressive disease and longer waiting times (namely, selection by time and LRT). This argument has been clearly shown also in "fast-track" living-donor LT and in studies about waiting time as a possible selector for the risk of post-transplant recurrence. (25-27) In all these studies, the patients presenting shorter waiting times also received a lower number of LRT.

Our study shows that performing a pre-transplant LRT strategy gives a beneficial effect on post-LT results as long as the number of treatments does not exceed the number of three. In other terms, when the number of required LRT for "stabilizing the tumor" is inferior or equal to three, the positive effect (namely, reducing the post-transplant recurrence) is statistically

significantly superior respect to the negative one (namely, increasing the de-listing rates). When four or more 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 LRT required, the higher the aggressiveness of the tumor, the worse the overall results.

Very low five-year HCC-dependent failure rates can be achieved in patients who initially received one LRT and did not exhibit disease progression after the treatment. Similar data were also observed in a large retrospective US analysis performed on 2,794 LT cases, in which a lower post-transplant recurrence rate was reported in patients undergoing LRT, while 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 post-transplant clinical course.

In an era in which great pressure exists on healthcare quality improvement and costs reduction, our study suggests the opportunity to frame a shift in standard practice towards LRT in MC-IN HCC patients. In our study, we observed a 34-49% risk reduction in patients receiving $\leq$3 LRT, 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 to 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 waiting time $\geq$ 6 months. On the opposite, studies clarifying the benefit of LRT in centres 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 **Additive Table 5**, we reported the different risks of de-listing, recurrence and overall HCC-dependent failure according to the use of an embolic vs. ablative vs. combined approach. Although a preliminary evidence exists on the fact that the necessity to treat the tumor with a combinatory approach during the waiting time should be connected with an increased risk of overall tumor-related failure, we should consider these preliminary results with caution. Further analyses focused 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 (i.e., LRT refractoriness of the target lesion vs. initial multimodal approach vs. 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 two decades and the period was introduced as a variable in the multivariable 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, trans-arterial radio-embolization or external radiotherapy should reduce the impact of our intention-to-treat analysis. On the opposite, we think that considering salvage LT after resection or very uncommon strategies like radio-embolization or radiotherapy in the present study should represent a bias, mainly due to their neglectable number in the present study.

Another possible limitation of the study concerns inter-center differences in the length of waiting time 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, since it enriches the variability within patients' 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 derives from the great initial difference among LRT and no-LRT cases make very difficult to construct a "balanced" model and that it is probably impossible to be eliminated in this specific setting.

CONCLUSIONS

Loco-regional therapy for Milan Criteria-IN HCC patients is valuable when considering an intention-to-treat liver transplant approach. When comparing treated and untreated patients, one single and two to three LRT confer a 49% and 34% reduction in the risk of HCC-dependent transplant failure. This beneficial effect disappears with more LRT, due to more aggressive tumor biology. Patients who show poor response to LRT have a predictably greater risk for pre-transplant tumor-related de-listing or post-transplant recurrence.

REFERENCES

1. Lai Q, Lerut JP. Hepatocellular cancer: how to expand safely inclusion criteria for liver transplantation. Curr Opin Organ Transplant 2014;19:229-34.

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;334:693-9.

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 Multicentre HCC Transplant Consortium. Ann Surg 2017;266:525-35.

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;21:248-57.

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;19:1343-53.

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;34:3661-79.

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; 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, Spoletini G, et al; 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, et al. mRECIST 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 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; 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;19:1108-18.

20. Duvoux C, Roudot-Thoraval F, Decaens T, Pessione F, Badran H, Piardi T, et al.; Liver Transplantation French Study Group. Liver transplantation for hepatocellular carcinoma: a model including α-fetoprotein improves the performance of Milan criteria. Gastroenterology 2012;143:986-94.

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 T-I, Huang Y-H, Su C-W, Lin H-C, Chiang J-H, Chiou Y-Y, 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, et al. 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;60:1957-62.

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;20:937-44.

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;57:974-9.

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.

APPENDIX

EurHeCaLT Study Group Collaborators

Austria: Konrad LEHNER (Medical University Innsbruck, Innsbruck);

Belgium: Maxime FOGUENNE, Isadora FRICK DE LA MAZA (Université Catholique de Louvain, Brussels);

Croatia: Iva KOSUTA (Merkur University Hospital, Zagreb);

Germany: Michael HEISE (University of Mainz, Mainz);

Italy: Enrico DALLA BONA (University of Padua, Padua); Fabio MELANDRO (Sapienza University, Rome); Leonardo BAIOCCHI (Tor Vergata University, Rome); Federico MOCCHEGIANI (Polytechnic University of Marche, Ancona); Giuseppe BIANCO (Catholic University of the Sacred Heart, Rome);

UK: Gabriele SPOLETINI (UCL Institute for Liver and Digestive Health and Royal Free Sheila Sherlock Liver Centre, Royal Free Hospital and UCL, London).

TABLES

Table 1. Patient-, tumor- and transplant-related characteristics of the pre-IPTW population.

Variables	No-LRT (n=182)	LRT (n=901)	p-value
	Median (IQR) or n (%)		
Age at HCC diagnosis §	58 (51-62)	58 (52-63)	0.70
Male gender §	153 (84.1)	741 (82.2)	0.59
Period of waiting list inscription 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			
max diameter target lesion §	2.0 (1.4-2.5)	2.4 (1.8-3.0)	<0.001
number lesions §	1 (1-2)	1 (1-2)	0.048
Radiological tumor features at LT or de-listing			
max diameter target lesion	2.0 (1.4-2.8)	1.7 (0.0-2.5)	<0.001
number 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 de-listing	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 de-listing			
CR	0 (-)	237 (26.3)	-
PR	0 (-)	258 (28.6)	-
SD	0 (-)	147 (16.3)	-
PD	0 (-)	259 (28.7)	-
Pathological tumor features **			
max diameter target lesion	2.0 (1.2-3.0)	2.0 (1.3-3.0)	0.40
number 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)	-

* Multiple pathologies in different patients; ** Calculated on 927 transplanted cases.
§ Variables used for performing the inverse probability treatment weighting.
Abbreviations: LRT, locoregional treatments; IQR, interquartile ranges; HCC, hepatocellular cancer; HCV, hepatitis C virus; HBV, hepatitis B virus; NASH, non-alcoholic steatohepatitis; MELD, model for end-stage liver disease; LT, liver transplantation; mRECIST, modified response evaluation criteria in solid tumors; CR, complete response; PR, partial response; SD, stable disease; PD, progression disease; MC, Milan Criteria; TACE, trans-arterial chemo-embolization; AFP, alpha-fetoprotein; WT, waiting time.

Table 2. Baseline characteristics of LRT and no-LRT subjects in the post-IPTW population.

Continuous variables	**Groups**	**Mean**	**SD**	**SE**	**Variance**	**SMD**	**ANOVA F-test**	**Student's t-test**
Age at listing	No-LRT	56.8	8.0	0.2	64.7	-3.7	0.73	0.73
	LRT	56.5	8.1	0.2	66.1			
Lab-MELD	No-LRT	13.5	5.0	0.2	25.4	-5.6	0.12	0.12
	LRT	13.2	5.7	0.2	32.9			
Major tumor diameter	No-LRT	2.4	1.0	0.03	1.0	0	0.76	0.76
	LRT	2.4	0.9	0.03	0.9			
Number of nodules	No-LRT	1.5	0.8	0.02	0.6	0	0.25	0.25
	LRT	1.5	0.7	0.02	0.5			
Log_{10}AFP ng/mL	No-LRT	1.9	2.7	1.2	5.4	0.6	0.19	0.19
	LRT	1.9	2.6	1.1	5.2			
Dichotomous variables	**Groups**	**Prevalence**			**SPD**		**Fisher's exact test**	
Male gender	No-LRT	84.6			-4.7		0.19	
	LRT	82.4						
HCV-positive status	No-LRT	46.6			-2.6		0.47	
	LRT	45.0						
HBV-positive status	No-LRT	17.4			1.3		0.73	
	LRT	18.0						

Abbreviations: SD, standard deviation; SE, standard error; SMD, standardized mean difference; ANOVA, analysis of variance; LRT, loco-regional treatment; MELD, model for end-stage liver disease; AFP, alpha-fetoprotein; HCV, hepatitis C virus; HBV, hepatitis B virus; SPD, standardized prevalence difference.

Table 3. Patient-, tumor- and transplant-related characteristics of the post-IPTW population.

Variables	No-LRT (n=1,077)	LRT (n=942)	p-value
	Median (IQR) or n (%)		
Age at HCC diagnosis §	58 (52-63)	58 (52-63)	0.73
Male gender §	911 (84.6)	776 (82.4)	0.19
Period of waiting list inscription 2001-2009	423 (54.8)	349 (45.2)	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			
max diameter target lesion §	2.2 (1.8-3.0)	2.3 (1.7-3.0)	0.75
number lesions §	1 (1-2)	1 (1-2)	0.25
Radiological tumor features at LT or de-listing			
max diameter target lesion	2.0 (1.6-3.0)	1.7 (0.0-2.4)	<0.001
number 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 de-listing	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 de-listing			
CR	0 (-)	253 (26.9)	-
PR	0 (-)	263 (27.9)	-
SD	0 (-)	147 (15.6)	-
PD	0 (-)	275 (29.2)	-
Pathological tumor features **			
max diameter target lesion	2.3 (1.4-3.0)	2.0 (1.3-3.0)	0.11
number 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)	-

* Multiple pathologies in different patients; ** Calculated on 1,791 transplanted cases.
§ Variables used for performing the inverse probability treatment weighting.
Abbreviations: LRT, locoregional treatments; IQR, interquartile ranges; HCC, hepatocellular cancer; HCV, hepatitis C virus; HBV, hepatitis B virus; NASH, non-alcoholic steato-hepatitis; MELD, model for end-stage liver disease; LT, liver transplantation; mRECIST, modified response evaluation criteria in solid tumors; CR, complete response; PR, partial response; SD, stable disease; PD, progression disease; MC, Milan Criteria; TACE, trans-arterial chemo-embolization; AFP, alpha-fetoprotein; WT, waiting time.

Table 4. Competing-risk model for the risk of HCC-dependent "LT strategy" failure in the pre- and post-IPTW populations.

Variables	Beta	SE	SHR	95%CI		p-value
				Inf	Sup	
PRE-IPTW						
WL inscription during the period 2001-2009	0.36	0.17	1.43	1.02	1.99	0.04
Male gender	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
Tumor major diameter at first referral	0.17	0.08	1.19	1.02	1.38	0.03
No LRT	Ref.	-	1.00	-	-	-
One LRT	-0.54	0.28	0.58	0.34	1.00	0.050
Two-three LRT	-0.32	0.27	0.73	0.43	1.23	0.24
Four or more LRT	-0.09	0.28	0.92	0.53	1.60	0.76
logAFP at first referral	0.42	0.10	1.53	1.26	1.86	<0.001
mRECIST progressive disease	1.74	0.17	5.70	4.11	7.90	<0.001
POST-IPTW						
WL inscription during the period 2001-2009	0.42	0.12	1.52	1.20	1.91	<0.001
Male gender	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
Tumor major diameter at first referral	0.27	0.05	1.31	1.18	1.45	<0.001
No LRT	Ref.	-	1.00	-	-	-
One LRT	-0.67	0.18	0.51	0.36	0.74	<0.001
Two-three LRT	-0.42	0.18	0.66	0.47	0.93	0.02
Four or more LRT	-0.22	0.20	0.80	0.55	1.17	0.27
logAFP at first referral	0.48	0.07	1.62	1.41	1.87	<0.001
mRECIST progressive disease	1.73	0.16	5.62	4.10	7.69	<0.001

Abbreviations: SE, standard error; SHR, semi-hazard ratios; CI, confidence intervals; IPTW, inverse propensity therapy weighting; WL, waiting list; MELD, model for end-stage liver disease; LRT, loco-regional treatment; AFP, alpha-fetoprotein; mRECIST, modified response evaluation criteria in solid tumors.

FIGURES

Figure 1A. HCC-dependent failure rates in the pre-IPTW population, stratified by number of LRTs.

Figure 1B. HCC-dependent failure rates in the post-IPTW population, stratified by number of LRTs.