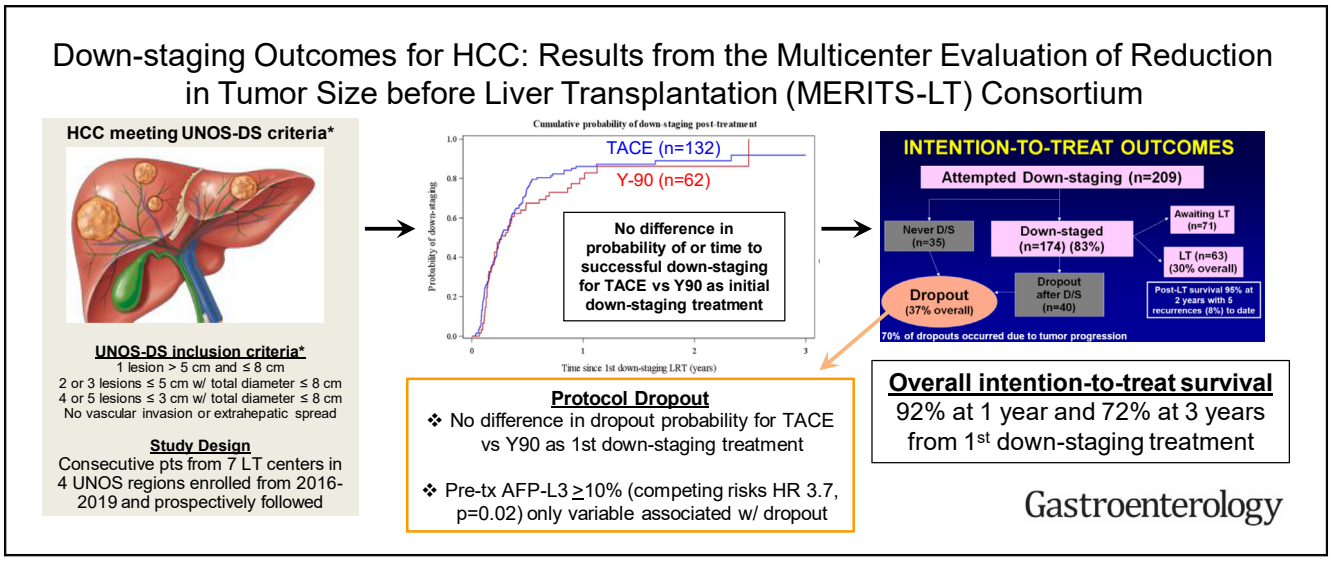
Gastroenterology 2021;161:1502–1512

Downstaging Outcomes for Hepatocellular Carcinoma: Results From the Multicenter Evaluation of Reduction in Tumor Size before Liver Transplantation (MERITS-LT) Consortium

Neil Mehta,[1] Catherine Frenette,[2] Parissa Tabrizian,[3] Maarouf Hoteit,[4] Jennifer Guy,[5] Neehar Parikh,[6] T. Tara Ghaziani,[7] Renu Dhanasekaran,[7] Jennifer L. Dodge,[8] Brahma Natarajan,[1] Matthew L. Holzner,[3] Leana Frankul,[2] Wesley Chan,[1] Austin Fobar,[6] Sander Florman,[3] and Francis Y. Yao[1,8]

[1]Division of Gastroenterology, Department of Medicine, University of California, San Francisco, California; [2]Center for Organ and Cell Transplantation, Scripps Green Hospital, La Jolla, California; [3]Department of Surgery, Icahn School of Medicine at Mount Sinai, New York, New York; [4]Division of Gastroenterology and Hepatology, University of Pennsylvania, Philadelphia, Pennsylvania; [5]Department of Transplantation, California Pacific Medical Center, San Francisco, California; [6]Division of Gastroenterology and Hepatology, University of Michigan, Ann Arbor, Michigan; [7]Division of Gastroenterology and Hepatology, Stanford University, Palo Alto, California; and [8]Division of Transplant Surgery, Department of Surgery, University of California, San Francisco, San Francisco, California

CLINICAL LIVER



BACKGROUND & AIMS: United Network of Organ Sharing (UNOS) has adopted uniform criteria for downstaging (UNOS-DS) of hepatocellular carcinoma (HCC) before liver transplantation (LT), but the downstaging success rate and intention-to-treat outcomes across broad geographic regions are unknown. **METHODS:** In this first multiregional study (7 centers, 4 UNOS regions), 209 consecutive patients with HCC undergoing downstaging based on UNOS-DS criteria were prospectively evaluated from 2016 to 2019. **RESULTS:** Probability of successful downstaging to Milan criteria and dropout at 2 years from the initial down-staging procedure was 87.7% and 37.3%, respectively. Pretreatment with lectin-reactive α-fetoprotein $\geq$10% (hazard ratio, 3.7; $P = .02$) was associated with increased dropout risk. When chemoembolization (n = 132) and yttrium-90 radioembolization (n = 62) were compared as the initial downstaging treatment, there were no differences in Modified Response Evaluation Criteria In Solid Tumors response, probability of or time to successful downstaging, waiting list dropout, or LT. Probability of LT at 3 years was 46.6% after a median of 17.2 months. In the explant, 17.5% had vascular invasion, and 42.8% exceeded Milan criteria (understaging). The only factor associated with understaging was the sum of the number of lesions plus largest tumor diameter on the last pre-LT imaging, and the odds of understaging increased by 35% per 1-unit increase in this sum. Post-LT survival at 2 years was 95%, and HCC recurrence occurred in 7.9%. **CONCLUSION:** In this first prospective multiregional study based on UNOS-DS criteria, we observed a successful downstaging rate of >80% and similar efficacy of chemoembolization and yttrium-90 radioembolization as the initial downstaging treatment. A high rate of tumor understaging was observed despite excellent 2-year post-LT survival of 95%. Additional LRT to reduce viable tumor burden may reduce tumor understaging.

Keywords: α-Fetoprotein (AFP); Local Regional Therapy (LRT); Tumor Recurrence; Waiting List Dropout.

Liver transplantation (LT) is an ideal treatment option for early-stage hepatocellular carcinoma (HCC) because LT removes not only the tumor but also the oncologic potential of the diseased liver. The number of HCC waiting list registrations in the United States has risen considerably in the past 2 decades, and HCC now accounts for nearly 30% of all LT performed in the United States.[1,2] The Milan criteria for LT[3] remain the gold standard for candidate selection in the United States, although many consider them to be too restrictive, and a plethora of expanded criteria have been proposed over the years.[4–7] The use of more liberal criteria, however, may result in higher tumor recurrence rates and reduce access to LT for other patients with a better prognosis.[8,9] Additionally, expanded criteria do not account for the effects of local regional therapy (LRT), which is increasingly used to control tumor growth when the waiting time is prolonged and also serves as a tool to improve candidate selection.[10] Regardless of whether the tumor stage is within or beyond Milan criteria, objective response to LRT has been shown to be a marker of favorable tumor biology, whereas tumor progression despite LRT reflects aggressive tumor behavior associated with a greater propensity for tumor recurrence after LT.[11–14]

Tumor downstaging, defined as a reduction of viable tumor burden by LRT to meet acceptable LT criteria, has garnered support in recent years as a better alternative to simply expanding the tumor size limits for LT.[6,7,15,16] In essence, downstaging aims at merging expanded criteria with response to LRT, serving as a tool to select a subset of patients with favorable tumor biology who would respond to downstaging treatments and also do well after LT.[15]

The University of California, San Francisco (UCSF) group published the largest single-center experience with tumor downstaging using a uniform protocol with well-defined inclusion criteria.[17] Those successfully downstaged to within Milan criteria had a 5-year post-LT survival of 78% and a tumor recurrence rate of 8%, similar to those initially meeting Milan criteria not requiring downstaging.[17] A subsequent study from United Network of Organ Sharing (UNOS) region 5, involving 3 centers[18] using the same downstaging protocol, showed similar results, with a 5-year post-LT survival of 80% and post-LT recurrence rate of <15%.

Despite these encouraging results, a pooled analysis of all published series on tumor downstaging demonstrated a wide range of downstaging success rates from 11% to 77% and tumor recurrence rates from 7 to 33%.[16] The substantial variations in these outcomes may be explained by the heterogeneity of the study populations and lack of strict inclusion criteria in most studies.

As many LT centers began to use tumor downstaging strategies for LT, staging definitions and end points varied widely across regions.[19] In an effort to standardize criteria for downstaging, the UNOS/Organ Procurement and Transplantation Network (OPTN) adopted the UCSF/region 5 downstaging protocol in 2017 (hereafter referred to as UNOS-DS) as a national policy, whereby patients meeting the UCSF/region 5 inclusion criteria and achieving successful downstaging to within Milan criteria are eligible to receive automatic priority listing with Model for End Stage Liver Disease (MELD) exception.[20] The standardized application of downstaging has also provided the opportunity for large-scale prospective multicenter studies to validate the feasibility and efficacy of tumor downstaging and to potentially refine selection or other staging criteria to further improve outcomes.[19]

In this first prospective multicenter study on downstaging from the Multicenter Evaluation of Reduction in Tumor Size before Liver Transplantation (MERITS-LT) consortium involving 7 centers from 4 UNOS regions, we aimed to examine the downstaging success rate and intention-to-treat outcomes based on uniform criteria (UNOS-DS protocol). We also sought to evaluate the

WHAT YOU NEED TO KNOW

BACKGROUND & CONTEXT

United Network of Organ Sharing downstaging protocol has been adopted for priority listing for liver transplant, although no national studies have confirmed the feasibility of downstaging or the optimal therapy to achieve successful downstaging.

NEW FINDINGS

Successful downstaging to within Milan criteria exceeded 80% with similar efficacy of transarterial chemoembolization and yittrium-90 as initial treatment. Although rates of explant understaging were relatively high, 2-year survival after liver transplant was 95%.

LIMITATIONS

There was short follow-up after liver transplant in a relatively small sample of liver transplant recipients and the possibility of referral bias.

IMPACT

This study validates the feasibility of down-staging on a broad scale under the current United Network of Organ Sharing downstaging guidelines.

CLINICAL LIVER

Abbreviations used in this paper: **AFP, α-fetoprotein; AFP-L3, lectin-reactive α-fetoprotein; CT, computed tomography; DCP, des-γ carboxyprothrombin; HCC, hepatocellular carcinoma; HR, hazard ratio; IQR, interquartile range; LRT, local regional therapy; LT, liver transplantation; MELD, Model for End Stage Liver Disease; mRECIST, modified Response Evaluation Criteria In Solid Tumors; MRI, magnetic resonance imaging; NLR, neutrophil-to-lymphocyte ratio; OPTN, Organ Procurement and Transplantation Network; OR, odds ratio; RETREAT, Risk Estimation of Tumor Recurrence After Transplant; TACE, transarterial chemoembolization; UCSF, University of California, San Francisco; UNOS, United Network for Organ Sharing; UNOS-DS, UNOS downstaging protocol; Y-90, yttrium-90.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2021.07.033

influence of the type of initial downstaging treatments and other factors on the likelihood of successful downstaging.

Materials and Methods

Downstaging Protocol and Radiographic Assessment

The UNOS-DS protocol has previously been described in detail[17] and is summarized in Table 1, including eligibility criteria based on initial tumor size and number and criteria for exclusion from LT. Consecutive patients from 7 high-volume LT centers with previous downstaging experience in 4 UNOS regions with HCC meeting UNOS-DS eligibility criteria were enrolled from 2016 to 2019 and prospectively monitored. Three additional LT centers from 2 more UNOS regions eventually were unable to provide data and were removed from the consortium. A minimum follow-up of 6 months after the first downstaging treatment was required for inclusion. All imaging studies for enrolled patients were assessed using Liver Imaging Reporting and Data System criteria,[21] which have been incorporated into the UNOS/OPTN guidelines.[22] Percutaneous biopsy was not routinely performed for the diagnosis of HCC at any of the institutions, and hepatic nodules <1 cm were not counted as HCC.

The specific type of LRT used was at the discretion of each of the center's multidisciplinary tumor boards and was not prespecified in the downstaging protocol. All patients included in the downstaging protocol underwent computed tomography (CT) or magnetic resonance imaging (MRI) of the abdomen 1 month after each LRT and at a minimum of once every 3 months. After LRT, radiographic assessment of tumor size was based on measurements of the maximum diameter of only viable tumors by multiphase CT or MRI and did not include the area of necrosis resulting from LRT.[23] Imaging criteria for successful downstaging included a decrease in tumor size to within Milan criteria or complete tumor necrosis with no contrast enhancement. Each center applied LRT with repetitive interventions if needed to achieve complete necrosis of all tumor nodules if possible. After successful downstaging of HCC, patients at each center were listed with MELD exception after a mandatory minimum waiting period of 6 months.

Histopathologic Analysis

In patients who underwent LT after successful downstaging, explant histopathologic features evaluated included tumor size, number of tumor nodules, histologic grade of differentiation,[24] and the presence or absence of microvascular or macrovascular invasion. Pathologic tumor staging was based on the UNOS TNM staging system.[25] The size and number of only viable tumors were considered in pathologic staging.

Outcomes

The primary outcome of interest was probability of and factors associated with successful downstaging and protocol dropout due to tumor progression or liver-related death, with the primary exposure being the type of the initial downstaging treatment. Secondary outcomes included probability of LT, post-LT survival, and HCC recurrence. For patients removed from the downstaging protocol for developing a medical contraindication to LT not related to liver disease, no longer

Table 1. United Network for Organ Sharing Downstaging Protocol

Inclusion criteria
HCC exceeding Milan criteria but meeting 1 of the following: 1. Single lesion 5.1–8 cm 2. 2–3 lesions each ≤5 cm with the sum of the maximal tumor diameters ≤8 cm 3. 4–5 lesions each ≤3 cm with the sum of the maximal tumor diameters ≤8 cm Plus absence of vascular invasion or extrahepatic disease based on cross-sectional imaging
Criteria for successful downstaging
Residual tumor size and diameter within Milan criteria (1 lesion ≤5 cm, 2–3 lesions ≤3 cm) a) Only viable tumor(s) are considered; tumor diameter measurements should not include the area of necrosis from tumor-directed therapy b) If there is more than 1 area of residual tumor enhancement, then the diameter of the entire lesion should be counted toward the overall tumor burden
Criteria for downstaging failure and exclusion from LT
1. Progression of tumor(s) to beyond inclusion/eligibility criteria for downstaging (as defined above) 2. Tumor invasion of a major hepatic vessel based on cross-sectional imaging 3. Lymph node involvement by tumor or extrahepatic spread of tumor 4. Infiltrative tumor growth pattern 5. Per current UNOS policy, if AFP ≥1000 ng/mL then transplant cannot be undertaken unless AFP level decreases to <500 ng/mL with LRT
Timing of LT in relation to downstaging
1. There should be a minimum observation period of 3 months of disease stability from successful downstaging to LT 2. Per current UNOS policy, the patient must remain within Milan criteria for 6 months after successful downstaging before receiving MELD exception points

interested in undergoing LT, or noncompliant with each center's transplant policies, follow-up was censored at the time of delisting or removal from the protocol.

Statistical Analysis

The date of the first downstaging procedure was defined as time 0 in all statistical analysis, except post-LT outcomes, for which the date of LT was time 0. Fisher's exact, χ^2, or Mann-Whitney U tests were used to compare differences in participant characteristics by type of first LRT (ie, transarterial chemoembolization [TACE] vs yttrium-90 [Y-90] radioembolization). The cumulative probabilities of successful downstaging, post-LT outcomes (survival and HCC recurrence), and intention-to-treat survival were estimated using the Kaplan-Meier method and compared across subgroups using the log-rank test. The cumulative probability of waiting list outcomes (dropout and LT) were estimated while accounting for the competing risk of the other waiting list event. Cox's regression assessed factors associated with successful downstaging were estimated as hazard ratios (HR). Fine and Gray competing-risks regression estimated the risk of dropout due to tumor progression or liver-related death as subHRs. Variables with a P value of <0.1 in univariable analysis were evaluated in bivariable models.

Results

Exclusion From the Downstaging Protocol

Among 324 patients with tumor burden meeting UNOS-DS criteria, 115 (35.5%) were not considered for LT and thus excluded from the downstaging protocol. Of these, 45.2% and 38.3% were due to medical and psychosocial contraindications to LT, respectively. Patients with baseline α-fetoprotein (AFP) >1000 ng/mL and decompensated liver disease have an exceedingly high risk of treatment failure and accounted for 13.0% of exclusions. Finally, 3.5% of exclusions were patients with a baseline bilirubin >4 mg/dL because they were not considered candidates for transarterial therapy.[17] The remaining 209 patients comprised the study cohort.

Baseline Characteristics and LRT

The baseline characteristics and details of LRT of the study cohort of 209 patients are presented in Table 2. Each center enrolled at least 12 patients, with the 2 largest centers enrolling 68 (32.5% of the overall cohort) and 33 patients (15.8%). The cohort was a median age of 63 years (interquartile range [IQR], 58–67 years), and 85.2% were men. White patients comprised 60.0% of the cohort, and hepatitis C was the most common etiology of liver disease (59.8%). At the time of the first downstaging procedure, median MELD was 9 (IQR, 7–11), 75.5% were Child's class A, and 3.0% were Child's class C. Median initial total tumor diameter was 6.2 cm (IQR, 5.6–7.3 cm). There were 32.1% with a single lesion measuring 5.1 to 8 cm, 54.1% with 2 to 3 lesions, and 13.9% with 4 to 5 lesions. Median baseline AFP was 13 ng/mL (IQR, 5–74 ng/mL), and 11.5% had a

Table 2. Baseline and Tumor Treatment Characteristics of the Downstaging Group

Study variable	Overall (N = 209)
Age, *y*	63 (58–67)
Male sex	178 (85.2)
Race/ethnicity	
White	123 (60.0)
Hispanic	45 (22.0)
Asian	23 (11.2)
African American	10 (4.9)
Liver disease etiology	
Hepatitis C	125 (59.8)
Alcohol	33 (15.8)
NAFLD	23 (11.0)
Hepatitis B	16 (7.7)
Other	12 (5.7)
CTP score[a]	6 (5–6)
Child's A (CTP 5–6)	151 (75.5)
Child's B (CTP 7–9)	43 (21.5)
Child's C (CTP 10–15)	6 (3.0)
MELD	9 (7–11)
AFP, *ng/mL*	13 (5–74)
>100 ng/mL	48 (23.0)
>1000 ng/mL	24 (11.5)
AFP-L3, %[b]	10.3 (4.6–16.9)
DCP[b]	2.5 (0.5–19.9)
Neutrophil-to-lymphocyte ratio	2.5 (1.7–3.8)
Platelet-to-lymphocyte ratio	86.3 (66.0–117.4)
No. of HCC lesions	
1 lesion	67 (32.1)
2–3 lesions	113 (54.1)
4–5 lesions	29 (13.9)
Initial total tumor diameter, *cm*	6.2 (5.6–7.3)
No. of LRT received	
1	44 (21.1)
2	53 (25.4)
3	41 (19.6)
4	25 (12.0)
>5	46 (22.0)
Type of LRT received	
Received 1 + TACE	169 (80.9)
Received 1 + Y-90	84 (40.2)
Received 1 + ablation	59 (28.2)
Type of first LRT received	
TACE	132 (63.1)
Y-90	62 (29.7)
Other	15 (7.2)

NOTE: Data are presented as median (IQR) or as n (%).
CTP, Child-Turcotte-Pugh; NAFLD, nonalcoholic fatty liver disease; No., number.
[a]n = 200.
[b]n = 83.

pretreatment AFP ≥1000 ng/mL. Median lectin-reactive α-fetoprotein (AFP-L3) was 10.3% (IQR, 4.6%-16.9%). Distribution of LRT received included 21.1% undergoing a single procedure, 25.4% receiving 2 LRTs, and 22.0% requiring ≥5 LRTs. TACE was the most common LRT used, with 80.9% receiving at least 1 such procedure, 40.2% receiving at least 1 Y-90 radioembolization, and 28.2% receiving at least 1 ablation procedure.

TACE was the first LRT received in 132 patients (63.1%), with 62 patients (29.7%) initially undergoing Y-90 tumor treatment. All centers performed both TACE and Y-90, although center-specific differences in the type of the first LRT were observed ($P = .001$). When baseline characteristics of these 2 first LRT groups (Table 3) were compared, median pretreatment age and Child's class were similar, but the TACE group had a higher proportion of men (92.4% vs 72.6%, $P < .001$) and slightly higher median MELD score ($P = .04$). The median total tumor diameter and pretreatment AFP were similar, but the TACE group was more likely to have multifocal disease (75.8% vs 51.6%, $P = .003$). Both groups had a median of 1 lesion (IQR, 1–2 lesions) treated with initial LRT.

Radiographic response by modified Response Evaluation Criteria In Solid Tumors (mRECIST) criteria to the first LRT

Table 3. Clinical Characteristics and Outcomes by Type of First Downstaging Treatment

Variable	TACE (n = 132)	Y-90 (n = 62)	*P* value
Age, *y*	63 (58–67)	63 (60–66)	.65
Male sex	122 (92.4)	45 (72.6)	<.001
CTP class[a]			.16
Child's A	91 (72.2)	50 (80.6)	
Child's B	29 (23.0)	12 (19.4)	
Child's C	6 (4.8)	0	
MELD	9 (7–12)	8.5 (7–10)	.04
AFP, *ng/mL*	11.7 (4.9–58.0)	17.9 (5.7–238.4)	.11
No. of HCC Lesions			.003
1 lesion	32 (24.2)	30 (48.4)	
2–3 lesions	80 (60.6)	25 (40.3)	
4–5 lesions	20 (15.2)	7 (11.3)	
Initial total tumor diameter, *cm*	6.3 (5.6–7.3)	6.3 (5.8–7.3)	.67
Lesions treated with first LRT, *No.*	1 (1–2)	1 (1–2)	.07
mRECIST response to first LRT			.67
Complete response	37 (28.0)	17 (27.4)	
Partial response	69 (52.3)	30 (48.4)	
Stable disease	14 (10.6)	7 (11.3)	
Progressive disease	12 (9.1)	8 (12.9)	
LRTs received, *No.*	3 (2–5)	2 (1–3)	.006
Ever downstaged	113 (85.6)	50 (80.6)	.38
Time to downstaged, *mo*	2.9 (1.3–5.6)	2.4 (1.7–4.6)	.73
Downstaging			
Protocol dropout	48 (36.4)	20 (32.3)	.58
Time to dropout, *mo*	8.4 (5.8–13.0)	10.2 (6.6–14.7)	.33
LT	44 (33.3)	14 (22.6)	.18
Time to LT, *mo*	18.3 (10.8–25.2)	15.9 (11.2–19.2)	.19
AFP before liver transplant, *ng/mL*	4.3 (3.0–21.7)	9.2 (6.0–16.0)	.18
Explant pathology			
Completely necrotic tumor(s)	9 (20.5)	4 (30.8)	.76
Beyond Milan	19 (43.2)	3 (23.1)	.44
Explant microvascular invasion	9 (20.5)	1 (7.7)	.29

NOTE: Data are presented as median (IQR) or n (%).
CTP, Child-Turcotte-Pugh; No., number.
[a]n = 188.

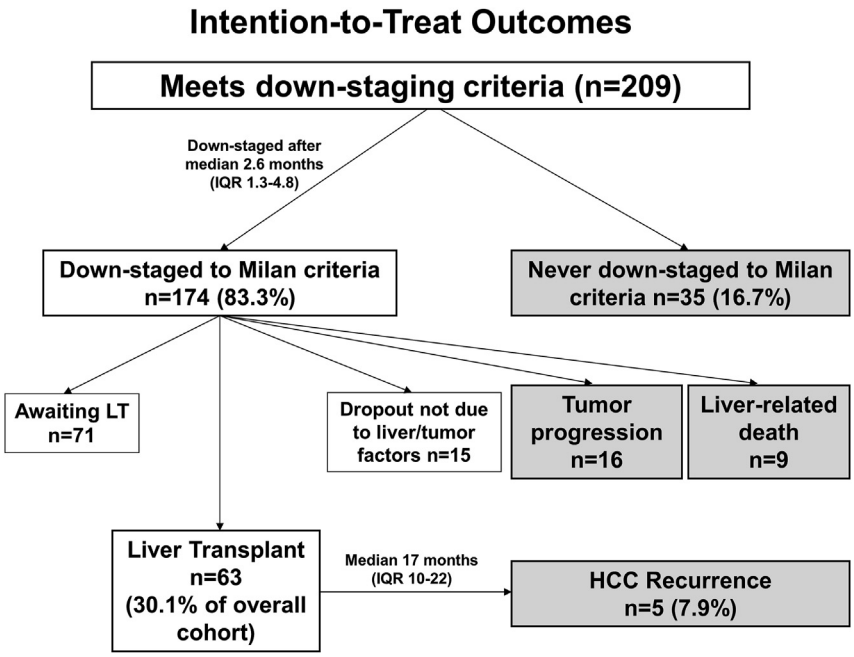


Figure 1. Summary of the intention-to-treat outcome of the 209 patients enrolled in the prospective downstaging protocol.

was similar between groups ($P = .67$), with partial response most common in both groups (TACE, 52.3%; Y-90, 48.4%), followed by complete response (TACE, 28.0%; Y-90, 27.4%). Median time from the initial LRT to posttreatment imaging on which the mRECIST response was assessed was slightly longer in the Y-90 group compared with the TACE group (6.3 vs 4.3 weeks, $P = .03$). The median number of total LRTs received was 3 (IQR, 2–5) in the TACE group compared with 2 (IQR, 1–3) in the Y-90 group ($P = .006$).

Intention-to-Treat Outcome

Tumor Downstaging. The intention-to-treat outcome related to attempted downstaging is summarized in Figure 1. Successful downstaging to within Milan criteria was achieved in 174 patients (83.3%) after a median of 2.6 months (IQR, 1.3–4.8 months). Among them, 66.1% were downstaged after a single LRT, whereas 33.9% required multiple treatments to achieve successful downstaging. The cumulative probability of successful downstaging to within Milan criteria from the first downstaging procedure was 67.5% at 6 months, 83.0% at 1 year, and 87.7% at 2 years. In Cox's regression models, the only factor associated with the ability to achieve tumor downstaging was decreasing total tumor burden (HR, 0.82 per cm; 95% CI, 0.69–0.96 per cm; $P = .02$). The probability of successful downstaging at 1 year from the first downstaging treatment was 88.3% in those with total tumor diameter <6 cm compared with 81.0% for total tumor diameter >7 cm ($P = .02$).

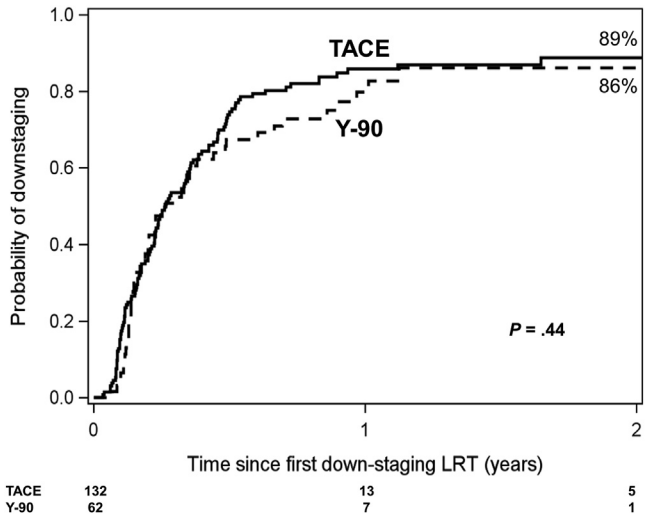


Figure 2. Kaplan-Meier probability of successful downstaging by the type of the first LRT with TACE vs Y-90.

When TACE and Y-90 as initial downstaging treatment were compared, no statistically significant differences were observed in probability of or time to successful downstaging (Table 3 and Figure 2). Additionally, pretreatment number of lesions, MELD score, Child's class, AFP, AFP-L3%, des-γ-carboxyprothrombin (DCP), and neutrophil-to-lymphocyte ratio (NLR) were not significant predictors of successful downstaging, nor was number of LRT received.

Downstaging Protocol Dropout. Of the 174 patients who initially achieved successful downstaging, 95.4% were subsequently listed for LT. Downstaging protocol dropout occurred in 75 patients (35.9% of overall cohort), including 51 due to tumor progression (68.0% of dropouts) and 9 (12.0%) due to liver-related death without LT. In those with tumor progression, 56.8% had dropout after receiving HCC MELD exception, and the remaining 43.2% were still in the initial 6-month waiting period at the time of dropout. The median time from the first downstaging treatment to dropout was 8.7 months (IQR, 5.9–13.4 months). The cumulative probability of dropout from the first downstaging procedure was 22.5% at 1 year and 37.3% at 2 years. In

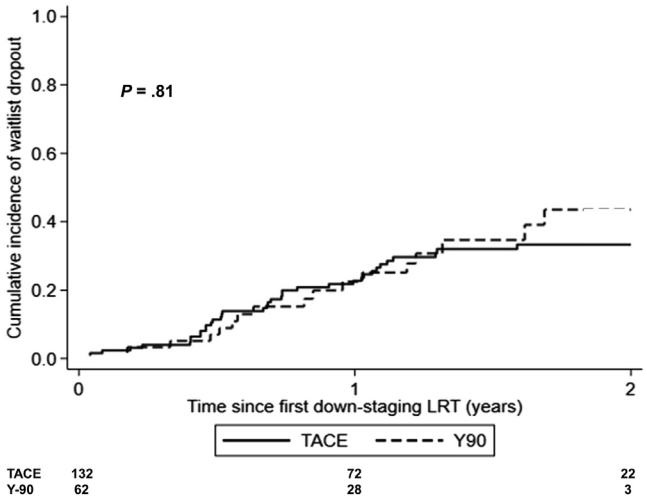


Figure 3. Kaplan-Meier probability of protocol dropout from the date of the first downstaging treatment.

bivariable competing-risks analysis, pretreatment with AFP-L3 ≥10% (subHR, 3.7; 95% CI, 1.27–10.79; $P = .02$) was associated with increased dropout due to tumor progression or liver-related death, even with separate adjustment for age or AFP. The probability of dropout within 3 years of the first LRT was 48.5% in those with an AFP ≥100 ng/mL compared with 37.3% for AFP <100 ng/mL ($P = .08$). No statistically significant differences were observed in probability of or time to dropout based on the type of the first LRT received (Table 3 and Figure 3), and no center-specific differences in dropout were observed.

Explant Pathology and Tumor Staging. At the last follow-up, 63 patients (30.1% of the entire cohort) had received LT, and 71 patients (34.0%) were within Milan criteria and active on the waiting list (Figure 1). On the last imaging before LT in these 63 patients, residual tumor was identified in 30 (47.6%), and the median sum of the largest viable lesion (cm) plus the number of viable lesions was 1.6 (IQR, 0–4.4). The median time from the first downstaging treatment to LT was 17.2 months (IQR, 11.1–24.3 months). The median time from successful downstaging to LT was 13.9 months and ranged from 9.6 months for the center with the shortest waiting time to 17.3 months for the center with the longest waiting time. The cumulative probability of LT at 1 and 3 years from the first downstaging procedure was 9.7% and 46.6%, respectively. When TACE and Y-90 were compared as the as initial downstaging treatment, no differences were observed in the proportion receiving LT, time to LT, or AFP at LT (Table 3) and no center-specific differences in probability of LT were observed.

At the time of LT, complete tumor necrosis from LRT (no residual tumors in explant) was observed in 23.8%. Tumor stage was within Milan criteria (T1/T2) in 33.3% and beyond Milan criteria (T3/T4) in 42.8% due to understaging by pre-LT imaging. The latter group included 1 patient with macrovascular invasion (T4b) and 1 with lymph node invasion (N1). Microvascular invasion was observed in 17.5%.

Among patients with viable tumors, 66.6% had moderately differentiated tumors, with 7 patients (14.6%) having poorly differentiated tumor grade. There were no significant differences in explant histology based on the type of first LRT received (Table 3), although Y-90 patients had a higher proportion with completely necrotic tumor(s) (30.8% vs 20.5%) and a lower proportion with both tumors beyond Milan criteria (23.1% vs 43.2%) and microvascular invasion (7.7% vs 20.5%; all $P > .25$). Overall, the median Risk Estimation of Tumor Recurrence After Transplant (RETREAT) score[25] was 2 (IQR, 1–3), with 8.1% having a RETREAT score of ≥5. The RETREAT score was similar based on type of first LRT received ($P = .56$).

On univariate logistic regression analysis, the only factor associated with explant understaging to beyond T2/Milan criteria was the sum of the number of lesions plus the largest tumor diameter on the last imaging before LT. The odds of understaging increased by 35% per 1-unit increase in this sum (odds ratio [OR], 1.35; 95% CI, 1.07–1.73; $P = .01$). Type of initial LRT, mRECIST response to initial LRT, type of last imaging before LT (ie, MRI vs CT), pre-LT AFP, and NLR both as continuous variables and at all tested cutoffs (AFP >20 and >100 ng/mL; NLR >5), and transplant center were not significant predictors of explant understaging on univariate analysis.

Factors associated with complete tumor necrosis in the explant using univariate logistic regression were pre-LT AFP <20 vs >20 ng/mL (OR, 11.6; $P = .007$) and the sum of the number of lesions plus largest tumor diameter on the last imaging before LT (OR, 0.72 per 1-unit increase; $P = .04$). Among those with no viable tumor on the last imaging before LT, 32.1% had complete tumor necrosis compared with 12.9% of those with suspected viable tumor on the last imaging (univariate logistic regression OR, 3.13; $P = .14$). Type of the initial LRT, mRECIST response to the initial LRT, type of last imaging before LT, pre-LT NLR, and transplant center were not associated with complete tumor necrosis in the explant.

Posttransplant Survival, Hepatocellular Carcinoma Recurrence, and Intention-to-Treat Survival. Among the 63 patients who underwent LT, median post-LT follow-up was 1.7 years (IQR, 1.2–2.4 years), and Kaplan-Meier post-LT survival at 1, 2, and 3 years was 100%, 95.0%, and 83.1%, respectively. HCC recurrence has developed in 5 patients (7.9%) to date, with a median time from LT to recurrence of 16.8 months (IQR, 9.7–22.3 months). In exploratory analysis, time from successful downstaging to LT was not associated with HCC recurrence. The overall Kaplan-Meier intention-to-treat survival at 1 and 3 years from the first downstaging procedure was 92.5% and 73.0%, respectively, with no significant difference found when TACE was compared with Y-90 as the type of the first LRT received. Stratified by initial tumor burden, intention-to-treat survival at 1 and 3 years from first LRT was 96.7% and 72.9%, respectively, in those with a total tumor diameter <6 cm compared with 91.3% and 72.6% for a total tumor diameter >7 cm ($P = .52$) (Figure 4).

Discussion

In this first multicenter prospective study on tumor downstaging from the MERITS-LT consortium (7 centers

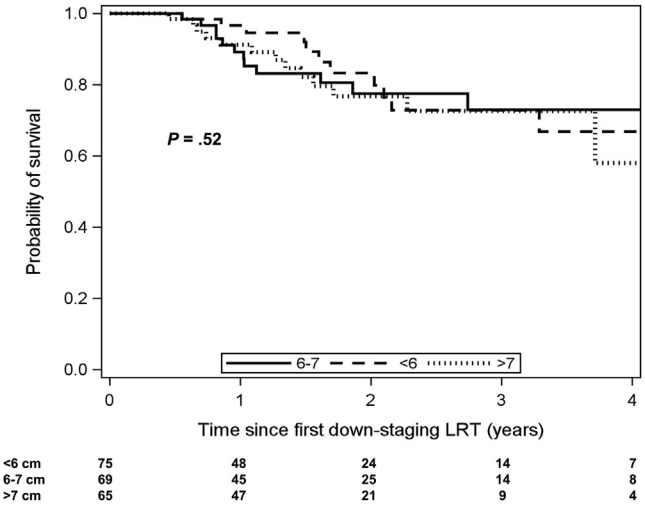


Figure 4. Kaplan-Meier probability of intention-to-treat survival from the first downstaging treatment stratified by the initial total tumor burden.

from 4 UNOS regions) designed to evaluate the outcomes of downstaging based on UNOS-DS criteria, we observed a very high overall probability of successful downstaging to within Milan criteria in 83% of the patients. The cumulative probabilities of successful downstaging were 68% at 6 months, 83% at 1 year, and 88% at 2 years after the first downstaging treatment. Approximately two-thirds were successfully downstaged after a single LRT. The only factor predicting successful downstaging was tumor burden measured by the sum of the largest tumor diameters. Even those with a total tumor diameter of >7 cm had an 81% probability of successful downstaging vs 88% for those with total tumor diameter of <6 cm.

These findings validate the feasibility of downstaging on a broad scale under the current UNOS-DS guidelines and highlight the importance of setting upper limits in the tumor burden to ensure a high downstaging success rate. Relaxing the eligibility criteria on initial tumor burden would result in significantly lower downstaging success rates[16,26,27] and potentially worse post-LT outcomes.[26,28] Sinha et al[26] reported an 84% rate of successful downstaging to Milan criteria in those meeting UCSF/UNOS-DS criteria, similar to the rate reported in the current study, vs a significantly lower success rate of 65% in the "all-comers" group with an initial tumor burden beyond these criteria and without upper limits. There was also a strong correlation between the sum of the tumor number and the largest tumor diameter and the likelihood of successful downstaging. The cumulative probability of successful downstaging at 1 year from the time of the first LRT decreased incrementally with a greater sum of the tumor number and largest tumor diameter and fell <50% in those with a sum of ≥12.[26]

A lesson learned from this and prior experience with downstaging is to restrict downstaging to only patients with adequate hepatic functional reserve. It has been proposed that only patients with Child's A or B cirrhosis with a total bilirubin ≤3 mg/dL should undergo tumor downstaging to ensure an acceptably low risk for posttreatment hepatic decompensation.[15] In the present study, we allowed enrollment of those with total bilirubin up to 4 mg/dL. In principle, patients, who develop hepatic decompensation after LRT before achieving successful downstaging are not eligible for LT. In the present study, 97% of patients had Child's A or B cirrhosis. The median Child-Pugh score was 6 and the median MELD score was 9 in our cohort. Improved selection of patients with good liver function might have contributed to the higher downstaging success rate in this study compared with previous reports on downstaging using the same tumor criteria for inclusion.[17,18]

It is important to point out that the present study followed a number of recently implemented UNOS guidelines, including a minimal waiting time of 6 months from successful downstaging to LT and exclusion of patients with AFP ≥1000 ng/mL from priority listing for LT unless there is a significant reduction of AFP to <500 ng/mL with LRT.[20] Under these guidelines, the cumulative probability of dropout due to tumor progression or liver-related death was 22.5% at 1 year and 37.3% at 2 years. Baseline AFP was not associated with the probability of successful downstaging or dropout, although there was a trend for an AFP ≥100 ng/mL to be associated with a higher risk of dropout. A baseline AFP >1000 ng/mL was not a predictor of a lower rate of successful downstaging or a higher risk of dropout, even though 12% of our cohort had a baseline AFP >1000 ng/mL and required a reduction to <500 ng/mL with LRT to be considered for LT.

In contrast, a previous study from UCSF[17] found a baseline AFP of ≥1000 ng/mL was associated with a 2.4–fold increased risk of waiting list dropout after downstaging. Similarly, a study from region 5 demonstrated an AFP of >1000 ng/mL was a significant predictor of treatment failure, defined as dropout, liver-related death without LT or HCC recurrence after LT.[18] The exclusion of Child's B or C patients with a baseline AFP >1000 ng/mL from enrollment into this study is a possible explanation for this discrepancy.

Although an AFP measurement is required in all patients at baseline and every 3 months while on the LT waiting list, other biomarkers, including AFP-L3, DCP, and NLR, were not obtained from all of the study population at baseline. Within these limitations, AFP-L3 ≥10% was the only factor independently associated with waiting list dropout. Although AFP-L3, DCP, and NLR have a potential role as prognostic markers in LT,[29–33] more prospective studies are needed to help define the place of these biomarkers in clinical practice.[34]

We also sought to assess the influence of the type of initial transarterial tumor treatment on downstaging outcomes. More than 90% of the cohort received TACE (n = 132) or Y-90 (n = 62) as the initial downstaging treatment. When we compared these 2 treatment modalities, pretreatment AFP and total tumor diameter were similar, and there were no observed differences in mRECIST response, probability of or time to successful downstaging, or probability of waiting list dropout or LT. Similarly, a systematic review and pooled analysis by Parikh et al[16] showed no difference in the rate of successful downstaging between Y-90 and TACE, although only 1 study compared these 2 treatment modalities.

In the present study, fewer LRTs were required for patients initially treated with Y-90 than those receiving TACE (median 2 [IQR 1–3] for Y-90 vs 3 [IQR, 2–5] for TACE). This finding mirrors that of the Prospective Randomized Study of Chemoembolization Versus Radioembolization for the Treatment of Hepatocellular Carcinoma. (PREMIERE) trial,[35] a small single-center phase II randomized trial that showed a significantly longer time to progression with Y-90 compared with TACE, but only 10 patients had tumors initially exceeding Milan criteria.

Although not statistically significant, we observed a higher rate of complete pathologic response and a lower probability of tumor understaging and microvascular invasion in the explant in those initially receiving Y-90. Until a large multicenter randomized trial comparing Y-90 and TACE is undertaken, the choice between these modalities as initial downstaging treatment will depend on center expertise and remain a matter of debate.

We observed excellent post-LT survival of 100% at 1 year and 95% at 2 years, but the follow-up was too short for post-LT outcomes to be the primary objective of this study. On a cautionary note, >40% had tumor understaging to beyond Milan criteria in the explant, which was at least 2-times higher than that in earlier studies from region 5[17,18] but in line with several recent analyses of the UNOS database.[28,36] In a study by Mehta et al,[28] one-third of patients with HCC initially meeting UNOS downstaging criteria had tumor beyond Milan criteria on explant. Another study by Mahmud et al[36] using the UNOS explant pathology form found tumor understaging in the explant was associated with increased post-LT HCC recurrence and death, and the risk of tumor understaging was higher among those requiring tumor downstaging before LT. Multiple explant-based prognostic models also demonstrated worse post-LT survival related to tumor understaging beyond Milan criteria in the explant.[32,37,38]

These findings underscore the importance of strict adherence to downstaging definitions and ensuring adequate response to downstaging before LT in addition to finding ways to reduce inaccuracies in radiographic staging assessments. It has been shown that in patients who require tumor downstaging, the higher the tumor burden on the last imaging study before LT, the greater the risk of understaging on explant pathology. The odds of tumor understaging on explant increases by 10% for each 1-cm increase in total tumor diameter on the last pre-LT imaging study.[28] Similarly, in the present study, the only factor associated with explant understaging to beyond Milan criteria was the sum of the number of lesions plus largest tumor diameter on the last imaging before LT. The odds of understaging increased by 35% per 1-unit increase in this sum. Based on these observations, we should consider downstaging to within Milan criteria as the minimal requirement for LT and perform additional LRT to further reduce the viable tumor burden and, ideally, to achieve complete tumor necrosis before LT.[19]

One of the strengths of this study is the prospective multicenter study design to investigate the outcomes of downstaging in a large cohort from 4 broad geographic regions using uniform inclusion criteria and end points of downstaging. Furthermore, the study period from 2016 to 2019 ensures no overlap of patients included in previous publications from several participating centers.[17,18]

There are also limitations, mainly the short duration of post-LT follow-up in a relatively small number of patients receiving LT to date. The primary objective of this study was to evaluate the feasibility of successful downstaging and waiting list outcomes. Another 2 to 3 years will be required to report long-term post-LT survival and HCC recurrence data to confirm the efficacy of tumor downstaging.

We originally sought to include 3 additional LT centers from 2 more UNOS regions to increase the study's power and applicability. However, these centers were unable to provide data and were therefore removed from the consortium.

This study was performed before the recent implementation of median MELD at LT minus 3 points for organ allocation for HCC. Consequently, this study could not account for the potential impact of such policy change on waiting list outcome and access to LT in downstaged patients. These effects should be analyzed in future studies.

Patients in all participating centers are eligible to receive both TACE and Y-90 radioembolization, although center-specific differences in the type of the first LRT for downstaging still exist. Specifically, the proportion undergoing TACE as the initial LRT (compared with Y-90) at the 7 centers ranges from 36% to 90%. This may be viewed as a real-world experience, and the type of initial LRT does not appear to have a significant impact on all the primary end points.

Finally, there is the possibility of a referral bias because patients within downstaging criteria who received LRT in the community but experienced subsequent disease progression might not have ever been referred for LT.

Conclusion

In this first prospective multiregional study based on UNOS-DS criteria, we observed a >80% probability of initial downstaging with relatively low likelihood of subsequent tumor progression and validated the feasibility of downstaging on a broad scale under the current UNOS-DS guidelines. We found similar efficacy of TACE and Y-90 as the initial downstaging treatment. Despite excellent 1- and 2-year post-LT survival, the tumor understaging rate was higher than expected. A point of emphasis is the critical importance of precise tumor staging definitions in achieving good outcomes.[19] Because pre-LT viable tumor burden strongly correlates with the risk of tumor understaging, we advocate downstaging to within Milan criteria as merely a minimal requirement for LT and performing additional LRT until complete tumor necrosis is achieved before LT.

References

1. Kwong A, Kim WR, Lake JR, et al. OPTN/SRTR 2018 annual data report: liver. Am J Transplant 2020;20(Suppl s1):193–299.

2. Mehta N, Dodge JL, Hirose R, et al. Increasing liver transplantation wait-list dropout for hepatocellular carcinoma with widening geographical disparities: implications for organ allocation. Liver Transpl 2018;24:1346–1356.
3. Mazzaferro V, Regalia E, Doci R, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N Engl J Med 1996; 334:693–699.
4. Yao FY. Liver transplantation for hepatocellular carcinoma: beyond the Milan criteria. Am J Transplant 2008; 8:1982–1989.
5. Sapisochin G, Bruix J. Liver transplantation for hepatocellular carcinoma: outcomes and novel surgical approaches. Nat Rev Gastroenterol Hepatol 2017;14:203–217.
6. Verna EC, Patel YA, Aggarwal A, et al. for the American Society of Transplantation Liver Intestinal Community of Practice. Liver transplantation for hepatocellular carcinoma: management after the transplant. Am J Transpl 2020;20:333–347.
7. Mehta N, Bhangui P, Yao FY, et al. Liver transplantation for hepatocellular carcinoma. Working group report from the ILTS Transplant Oncology Consensus Conference. Transplantation 2020;104:1136–1142.
8. Clavien PA, Lesurtel M, Bossuyt PM, et al; OLT for HCC Consensus Group. Recommendations for liver transplantation for hepatocellular carcinoma: an international consensus conference report. Lancet Oncol 2012; 13:e11–e22.
9. Volk ML, Vijan S, Marrero JA. A novel model measuring the harm of transplanting hepatocellular carcinoma exceeding Milan criteria. Am J Transplant 2008;8:839–846.
10. Mehta N, Yao FY. Moving past "one size (and number) fits all" in the selection of candidates with hepatocellular carcinoma for liver transplant. Liver Transpl 2013; 19:1055–1058.
11. Lai Q, Avolio AW, Graziadei I, et al. 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.
12. Otto G, Herber S, Heise M, et al. Response to transarterial chemoembolization as a biological selection criterion for liver transplantation in hepatocellular carcinoma. Liver Transpl 2006;12:1260–1267.
13. Kim DJ, Clark PJ, Heimbach J, et al. Recurrence of hepatocellular carcinoma: importance of mRECIST response to chemoembolization and tumor size. Am J Transplant 2014;14:1383–1390.
14. Mehta N, Dodge JL, Goel A, et al. 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.
15. Yao FY, Fidelman N. Reassessing the boundaries of liver transplantation for hepatocellular carcinoma: where do we stand with tumor down-staging? Hepatology 2016; 63:1014–1025.
16. Parikh ND, Waljee AK, Singal AG. Downstaging hepatocellular carcinoma: a systematic review and pooled analysis. Liver Transpl 2015;21:1142–1152.
17. Yao FY, Mehta N, Flemming J, et al. Downstaging of hepatocellular cancer before liver transplant: long-term outcome compared to tumors within Milan criteria. Hepatology 2015;61:1968–1977.
18. Mehta N, Guy J, Frenette CT, et al. Excellent outcomes of liver transplantation following down-staging of hepatocellular carcinoma to within Milan criteria: a multicenter study. Clin Gastroenterol Hepatol 2018; 16:955–964.
19. Yao F, Fidelman N, Mehta N. The key role of staging definitions for assessment of downstaging for hepatocellular carcinoma. Semin Liv Dis 2021;41:117–127.
20. Organ Procurement and Transplantation Network. Changes to HCC Criteria for Auto Approval; 2017. United States Department of Health and Human Services. https://optn.transplant.hrsa.gov/media/1922/liver_hcc_criteria_for_auto_approval_20160815.pdf. Accessed September 20, 2020.
21. Mitchell DG, Bruix J, Sherman M, et al. LI-RADS (Liver Imaging Reporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions. Hepatology 2015; 61:1056–1065.
22. Wald C, Russo MW, Heimbach JK, et al. New OPTN/UNOS policy for liver transplant allocation: standardization of liver imaging, diagnosis, classification, and reporting of hepatocellular carcinoma. Radiology 2013; 266:376–382.
23. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis 2010;30:52–60.
24. Edmondson HA, Steiner PE. Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. Cancer 1954;7:462–503.
25. Pomfret EA, Washburn K, Wald C, et al. Report of a national conference on liver allocation in patients with hepatocellular carcinoma in the United States. Liver Transpl 2010;16:262–278.
26. Sinha J, Mehta N, Dodge JL, et al. Are there upper limits in tumor burden for down-staging of HCC to liver transplant? Analysis of the all-comers protocol. Hepatology 2019;70:1185–1196.
27. Murali AR, Romero-Marrero C, Miller C, et al. Predictors of successful downstaging of hepatocellular carcinoma outside Milan criteria. Transplantation 2016;100:2391–2397.
28. Mehta N, Dodge JL, Grab JD, et al. National experience on down-staging of hepatocellular carcinoma before liver transplant: influence of tumor burden, AFP, and wait time. Hepatology 2020;71:943–954.
29. Chaiteerakij R, Zhang X, Addissie BD, et al. Combinations of biomarkers and Milan criteria for predicting hepatocellular carcinoma recurrence after liver transplantation. Liver Transpl 2015;21:599–606.
30. Lee JH, Cho Y, Kim HY, et al. Serum tumor markers provide refined prognostication in selecting liver transplantation candidate for hepatocellular carcinoma

patients beyond the Milan criteria. Ann Surg 2016; 263:842–850.
31. Fujiki M, Takada Y, Ogura Y, et al. Significance of des-gamma-carboxy prothrombin in selection criteria for living donor liver transplantation for hepatocellular carcinoma. Am J Transpl 2009;9:2362–2371.
32. Halazun KJ, Najjar M, Abdelmessih RM, et al. Recurrence after liver transplantation for hepatocellular carcinoma: a new MORAL to the story. Ann Surg 2017; 265:557–564.
33. Lai Q, Castro Santa E, Rico Juri JM, et al. Neutrophil and platelet-to-lymphocyte ratio as new predictors of dropout and recurrence after liver transplantation for hepatocellular cancer. Transpl Int 2014;27:32–41.
34. Kotwani P, Chan WO, Yao F, et al. DCP and AFP-L3 are complementary to AFP in predicting high-risk explant features: results of a prospective study. Clin Gastroenterol Hepatol. Published online January 29. https://doi.org/10.1016/j.cgh.2021.01.043.
35. Salem R, Gordon AC, Mouli S, et al. Y90 radioembolization significantly prolongs time to progression compared with chemoembolization in patients with hepatocellular carcinoma. Gastroenterology 2016; 151:1155–1163.
36. Mahmud N, Hoteit MA, Goldberg DS. Risk factors and center-level variation in hepatocellular carcinoma understaging for liver transplantation. Liver Transpl 2020; 26:977–988.
37. Agopian VG, Harlander-Locke M, Zarrinpar A, et al. A novel prognostic nomogram accurately predicts hepatocellular carcinoma recurrence after liver transplantation: analysis of 865 consecutive liver transplant recipients. J Am Coll Surg 2015;220:416–427.
38. Mehta N, Heimbach J, Harnois DM, et al. Validation of a Risk Estimation of Tumor Recurrence After Transplant (RETREAT) score for hepatocellular carcinoma recurrence after liver transplant. JAMA Oncol 2017;3:493–500.

Received January 21, 2021. Accepted July 20, 2021.

Correspondence
Address correspondence to: Neil Mehta, MD, University of California, San Francisco, 513 Parnassus Avenue, Room S-357, San Francisco, California 94143-0538. e-mail: neil.mehta@ucsf.edu.

CRediT Authorship Contributions
Neil Mehta, MD (Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Investigation: Lead; Methodology: Lead; Writing – original draft: Lead; Writing – review & editing: Lead). Catherine Frenette, MD (Data curation: Supporting; Writing – review & editing: Supporting). Parissa Tabrizian, MD (Data curation: Supporting; Writing – review & editing: Supporting). Maarouf Hoteit, MD (Data curation: Supporting; Writing – review & editing: Supporting). Jennifer Guy, MD (Data curation: Supporting; Writing – review & editing: Supporting). Neehar Parikh, MD (Data curation: Supporting; Writing – review & editing: Supporting). T. Tara Ghaziani, MD (Data curation: Supporting; Writing – review & editing: Supporting). Renu Dhanasekaran, MD (Data curation: Supporting; Writing – review & editing: Supporting). Jennifer L. Dodge, MPH (Data curation: Supporting; Formal analysis: Lead; Methodology: Lead; Writing – review & editing: Supporting). Brahma Natarajan, MD (Data curation: Supporting). Matthew L. Holzner, MD (Data curation: Supporting; Writing – review & editing: Supporting). Leana Frankul, BS (Data curation: Supporting). Wesley Chan, BS (Data curation: Supporting; Writing – review & editing: Supporting). Austin Fobar, BS (Data curation: Supporting). Sander Florman, MD (Supervision: Supporting; Writing – review & editing: Supporting). Francis Y. Yao, MD (Conceptualization: Supporting; Supervision: Lead; Writing – review & editing: Equal).

Conflicts of interest
These authors disclose the following: Neil Mehta has served on advisory boards for Wako Diagnostics and has received institutional research funding from Wako Diagnostics, Glycotest, and Target PharmaSolutions. Maarouf Hoteit has served on advisory boards for Eisai and HepQuant. Neehar Parikh serves as a consultant for Bristol Myers-Squibb, Exact Sciences, Eli Lilly, and Freenome, has served on advisory boards for Genentech, Eisai, Bayer, Exelixis, and Wako Diagnostics, and has received institutional research funding from Bayer, Target PharmaSolutions, Exact Sciences, and Glycotest. Francis Yao has received institutional research funding from Wako Diagnostics. The remaining authors disclose no conflicts.

Funding
This work is supported in part by a grant from the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases to the University of California, San Francisco Liver Center (P30 DK026743).