Liver Transplantation for Advanced Hepatocellular Carcinoma after Downstaging Without Up-Front Stage Restrictions

William C Chapman, MD, FACS, Sandra Garcia-Aroz, MD, Neeta Vachharajani, BS, Kathryn Fowler, MD, Nael Saad, MD, Yiing Lin, MD, PhD, Jason Wellen, MD, FACS, Benjamin Tan, MD, Adeel S Khan, MD, FACS, MB Majella Doyle, MD, FACS

BACKGROUND: The incidence of hepatocellular carcinoma (HCC) continues to increase dramatically worldwide. Liver transplantation (LT) is now the standard and optimal treatment for patients with HCC in the setting of cirrhosis, but only for tumors within Milan criteria. In patients presenting beyond Milan criteria, locoregional therapy (LRT) can downstage to within Milan criteria for consideration for LT. Although controversial, the current study aims to evaluate the outcomes of LT in patients presenting with advanced-stage HCC who underwent downstaging and compare these outcomes with those of patients who met Milan criteria at presentation.

STUDY DESIGN: Our protocol does not set a priori limitations as long as HCC is confined to the liver. In this retrospective study between January 1, 2002 and December 31, 2014, we reviewed outcomes associated with 284 patients who presented within Milan criteria and patients who presented with more-advanced stage tumor who were potential transplantation candidates. The patients with advanced disease were then subdivided into those who were within or beyond University of California San Francisco criteria. Imaging, details of LRT, recurrence, and survival were compared between the groups.

RESULTS: Sixty-three of 210 (30%) eligible patients were downstaged and underwent transplantation; 14 additional downstaged and listed patients were withdrawn for the following reasons: death while waiting (n = 4), disease progression (n = 8), development of other malignancy (n = 1), and declined LT (n = 1). Twelve patients underwent resection after downstaging and did not require LT. Survival for patients who were downstaged was similar to those who were within Milan criteria initially. Recurrence of HCC at 5 years was similar between groups (10.9% vs 10.8%; p = 0.84).

CONCLUSIONS: Patients with beyond-Milan criteria HCC who are otherwise candidates for LT should undergo aggressive attempts at downstaging without a priori exclusion. This highly selective approach allows for excellent long-term results, similar to patients presenting with earlier-stage disease. (J Am Coll Surg 2017;224:610–621. Published by Elsevier Inc. on behalf of the American College of Surgeons.)

Hepatocellular carcinoma (HCC) is the fifth most common malignancy worldwide and the second leading cause of cancer-related mortality in the world, responsible for about 1 million deaths per year. In the US, HCC is predicted to become one of the leading causes of cancer-related mortality by 2030. Noninvasive criteria for HCC diagnosis are based on imaging techniques obtained by 4-phase multidetector CT scan or dynamic

Disclosure Information: Nothing to disclose.
Presented at the Southern Surgical Association 128th Annual Meeting, Palm Beach, FL, December 2016.

Received December 19, 2016; Accepted December 19, 2016.
From the Departments of Surgery (Chapman, Garcia-Aroz, Vachharajani, Lin, Wellen, Khan, Doyle), Radiology (Fowler, Saad), and Medical Oncology (Tan), Washington University School of Medicine, St Louis, MO.
Correspondence address: William C Chapman, MD, FACS, Department of Surgery, Washington University School of Medicine, 660 S Euclid Ave, Campus Box 8109, St Louis, MO 63110. email: chapmanw@wustl.edu

Published by Elsevier Inc. on behalf of the American College of Surgeons.
Hepatocellular Carcinoma Downstaging

Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFP</td>
<td>α-fetoprotein</td>
</tr>
<tr>
<td>ALTS TG</td>
<td>American Liver Tumor Study Group</td>
</tr>
<tr>
<td>B-</td>
<td>beyond University of California San Francisco</td>
</tr>
<tr>
<td>UCSF</td>
<td>criteria</td>
</tr>
<tr>
<td>HCC</td>
<td>hepatocellular carcinoma</td>
</tr>
<tr>
<td>LRT</td>
<td>locoregional therapy</td>
</tr>
<tr>
<td>LT</td>
<td>liver transplantation</td>
</tr>
<tr>
<td>MELD</td>
<td>Model for End-Stage Liver Disease</td>
</tr>
<tr>
<td>RFA</td>
<td>radiofrequency ablation</td>
</tr>
<tr>
<td>TACE</td>
<td>transarterial chemoembolization</td>
</tr>
<tr>
<td>UCSF</td>
<td>University of California San Francisco</td>
</tr>
<tr>
<td>Y-90</td>
<td>yttrium-90 radioembolization</td>
</tr>
</tbody>
</table>

Locoregional therapy includes transarterial chemoembolization (TACE), selective yttrium-90 radioembolization (Y-90), and/or radiofrequency ablation (RFA). These techniques, when used in the setting of multidisciplinary care and in properly selected patients, can be used to downstage tumor burden to within an acceptable range, that is, within Milan criteria. The consideration of more-advanced stage patients, especially those who are downstaged using LRT, for LT remains controversial and additional validation of good outcomes is necessary, given the limited pool of transplantable organs. In this study, we present 63 patients with HCC beyond Milan criteria who underwent LT after successful downstaging to within Milan criteria. We compared the results with patients, initially within Milan criteria, who received transplants in the same time period.

METHODS

Approval was obtained from the Washington University IRB for this retrospective study. Our prospective liver transplant database was queried to identify HCC patients who received transplants between January 1, 2002 and December 31, 2014. Their original HCC disease extent was analyzed based on Milan criteria and the American Liver Tumor Study Group (ALTS TG) staging. Variations potentially impacting survival and disease recurrence after LT were identified and compared.

Throughout the majority of the study timeline, HCC tumors were diagnosed by cross-sectional imaging according to the diagnostic criteria defined in the American Association for the Study of Liver Diseases guidelines, which require a contrast-enhancing mass >1 cm in size in the arterial phase with venous or delayed phase washout in the setting of cirrhosis. After implementation of a new national policy on October 31, 2013, the more stringent Organ Procurement and Transplantation Network criteria for imaging diagnosis were used. Liver biopsies were not performed unless atypical features were present.

Inclusion criteria for the study were age older than 18 years at HCC diagnosis; a single nodule >5 cm, 2 to 3 nodules at least 1 >3 cm, corresponding to stage III of the ALTS TG Classification, or ≥4 nodules of any size (stage IVa1 of the ALTS TG classification), or HCC with any tumor stage plus intrahepatic portal or hepatic vein involvement (stage IVa2 of the ALTS TG classification); possible transplantation candidate; and able to undergo LRT for possible downstaging. α-Fetoprotein (AFP) level was not included as an inclusion or exclusion criteria. Patients with regional lymph nodes or metastatic disease (including extrahepatic main portal or hepatic vein involvement),
stage IVB of ALTSG classification, were excluded from this study. Two hundred and ten patients met the inclusion criteria and followed the downstaging protocol.16

Locoregional therapy and evaluation of response

Each patient was evaluated by a multidisciplinary team, and TACE, Y-90, RFA, or a combination of these therapies, were applied according to case-by-case discussion. In general, patients with unilobar disease were treated with a single TACE session. The technique and procedure have been described previously.14 Patients with bilobar disease were treated at 4- to 6-week intervals with serial TACE until all nodules were treated or treatment failure occurred due to progression of underlying liver disease precluding additional therapy. When the treatment was completed, a new imaging study (contrasted CT or MRI) was performed to evaluate the liver 4 to 6 weeks after the last LRT. Viability of tumor on follow-up imaging studies was determined by the presence of arterial phase hyper-enhancement within the nodule, as described by studies was determined by the presence of arterial phase hyper-enhancement within the nodule, as described by modified Response Evaluation Criteria in Solid Tumors. Assessment of tumor response took into account treated lesions, untreated lesions, and any new observations on follow-up imaging. Consideration for transplantation was made when residual tumor burden after treatment met ALTSG stage II criteria (1 nodule <5 cm or 2 to 3 nodules, all <3 cm). The presence of extrahepatic disease, progression of intrahepatic disease, or failure to achieve stage II criteria resulted in exclusion from consideration for LT. All patients were submitted to the United Network for Organ Sharing Regional Review Board for Region 8 after downstaging was completed and MELD exception points were granted, similar to patients initially within Milan criteria.

The number of LRTs and time needed for downstaging were recorded. Tumor progression, other malignancies, and death were also noted. Tumor burden at baseline was described using Milan criteria and UCSF criteria.11,17 Patients were classified as beyond Milan criteria/within UCSF (W-UCSF), or beyond UCSF (B-UCSF) at baseline. Success of downstaging and LT rate was assessed for each group. LT outcomes (overall survival, graft survival, disease-specific survival, disease-free survival, and recurrence rate) were also analyzed.

Control group

To compare results of our study group (patients beyond Milan criteria), we also collected data of all patients with HCC who met Milan criteria at the outset and underwent LT between January 2002 and December 2014. Demographic, clinical, and tumor characteristics were studied. Outcomes after LT (overall survival, graft survival, disease-specific survival, disease-free survival, and recurrence rate) were also analyzed for this group.

Statistical analysis

Statistical analyses were done using GraphPad Prism, version 5.0 (GraphPad Software). Categorical variables, presented as n (%), were compared using Fisher’s exact test or chi-square test, as applicable. Continuous variables presented as median (range) or mean (SD) were compared using one-way ANOVA (for 3-group comparison) or Mann-Whitney test (for intention-to-treat analysis). Survival graphs were plotted using Kaplan-Meier curves and survival rates were compared using log-rank test. Multivariate analysis was run using SPSS, version 23 (IBM Corp). For every comparison, differences with a p value <0.05 were considered statistically significant.

RESULTS

Between January 1, 2002 and December 31, 2014, four hundred and fifty-six new patients presenting with beyond-Milan criteria HCC without metastatic disease received LRT at our center. Of these, 246 patients were not considered future transplantation candidates from presentation; 28 were aged older than 80 years, 28 were not surgical candidates, 27 had social contraindications, and 47 had major comorbidities. Within 3 months from their last LRT, 78 patients died and 38 were lost to follow-up. The remaining 210 patients with beyond-Milan criteria HCC were considered possible future LT candidates at the time, initiating the downstaging LRT protocol.

Sixty-three of 210 (30%) eligible patients were downstaged and received transplants; 14 (6.7%) additional downstaged and listed patients were withdrawn for the following reasons: death while waiting (n = 4), disease progression (n = 8), development of other malignancy (n = 1), and declined LT (n = 1). Twelve (5.7%) underwent resection after downstaging and did not require LT. One hundred and twenty-one patients began LRT with an intention to downstage to within Milan criteria, but failed to reach it (Fig. 1).

Demographics and tumor characteristics

Demographics, clinical parameters, and tumor characteristics of the study population are listed in Table 1. Mean ± SD age at first LRT was 61.5 ± 10.4 years. Mean ± SD AFP level at the time of first LRT was 5,788 ± 21,728 ng/mL. The most frequent underlying liver disease was hepatitis C virus (44.3%), followed by alcoholic liver disease (9.5%). Thirty-five patients (16.7%) were W-UCSF and 175 patients (83.3%) were B-UCSF. The most commonly used LRT was TACE, either alone (n = 171 [81.4%]) or in combination with Y-90.
The remaining 23 (11.0%) patients received Y-90 alone. Eleven patients (5.2%) received RFA in addition to either TACE (n = 10) or Y-90 (n = 1). The number of LRTs received was also recorded. Patients received a median of 2 (range 1 to 6) LRTs.

Of these patients, 89 were successfully downstaged. Demographic and tumor characteristics were compared to assess the differences between groups (downstaged vs not downstaged after LRT). Results are also described in Table 1. Significant differences were found when we compared age at first LRT, 58 vs 63 years old, respectively (p = 0.004). Presence of hepatitis C virus cirrhosis was significantly higher in the downstaged group (56.2% vs 35.3%; p = 0.003). The size of the largest tumor was significantly greater in the patients who did not achieve downstaging (6.5 vs 5.0 cm; p = 0.002). Numerical differences were present for the ALTSG tumor stage and number of tumors at presentation, with a higher percentage of patients with more advanced stage and greater number of tumors not achieving downstaging, although these differences did not achieve statistical significance.

Locoregional therapy response

Downstaged group

Thirty-five patients showed a complete response by modified Response Evaluation Criteria in Solid Tumors, and 52 patients presented a partial response with viable tumor demonstrated in the CT/MRI pretransplantation. Two patients showed stable disease.

Fourteen patients were listed but were later withdrawn; 3 patients had complete response and 11 had partial response. Of the 12 patients that received resection, 2 demonstrated pathologic complete response, 8 had pathologic partial response, and 2 showed stable disease.

For the 63 patients that underwent LT after downstaging, a comparison between the radiologic description and the pathology results was performed. Thirty of these 63 patients achieved complete response to LRT based on cross-sectional imaging before transplantation. Fourteen of these did not show viable tumor on explant, 8 showed <20% tumor viability, and 8 showed >20% viability. A total of 9 patients showed beyond-Milan criteria disease in the explant. Seven of these 9 cases were beyond Milan criteria because of the multiple sub-centimeter tumors. The remaining 2 explants showed vascular thrombus, which was not seen on pre-LT imaging.

Failure to achieve downstaging group

Of the 121 patients that failed to achieve downstaging to Milan criteria, most patients showed partial response (n = 58 [47.9%]), but disease progressed in 28 of these 58 patients in short-term follow-up. Of these 28 patients that showed partial response followed by disease progression, 4 patients received follow-up LRT (RFA [n = 3] and TACE [n = 1]) as palliative care, 4 patients received sorafenib therapy, 7 died within 3 months from that imaging, 7 were lost to follow-up after that imaging, and 6 patients chose the no-treatment option. Of the 6 that did not accept the offered treatment options, 2 died 4 and 23 months later, 3 were lost to follow-up with extensive disease 13, 14, and 46 months later, and 1 still follows with us (24 months) and disease has not progressed. Of the remaining 30 patients in whom disease progression was not documented, 21 additional patients died (14 within 3 months after showing partial response, 4 within 3 to 6 months, and 3 within 6 to 12 months). Of the remaining 9 of these patients, 3 were lost to follow-up 3 months after last imaging, 2 received additional LRTs as palliative care, and the others were not eligible or not interested in any definitive treatment.

Factors affecting downstaging success

Binary logistic regression analysis was performed to assess the effect of age at first LRT (age older than 70 years), presence of cirrhosis, hepatitis C virus liver disease,
Y-90 therapy, >3 LRTs given, >3 tumors at presentation, largest tumor diameter >8 cm at presentation, and AFP level >1,000 ng/mL at first LRT on the downstaging outcomes. The analysis showed that younger (70 years or younger at first LRT) patients with cirrhosis, with ≤3 tumors, and no tumor thrombus at the time of disease identification were more likely to downstage. Contrary to our expectation, AFP >1,000 ng/mL at presentation did not predict failure of successful downstaging.

**Long-term outcomes in hepatocellular carcinoma patients**

As described previously, between January 2002 and December 2014, 63 patients with HCC were beyond...
Milan criteria at diagnosis and, after successful downstaging, underwent LT. Of these 63 patients, 18 were W-UCSF, and 45 patients were beyond Milan criteria and B-UCSF criteria.

During the same period, 289 transplantations were performed at our center in patients that presented HCC within Milan criteria when the disease was identified. Fifty of these 289 patients were identified incidentally in the liver explant after LT, and were excluded from this analysis. Therefore, 239 cases of HCC patients within Milan criteria at diagnosis represent our control group (within Milan criteria) for this study.

Demographics

Results of univariate analysis comparing the within Milan criteria, W-UCSF, and B-UCSF groups are presented in Table 2. A higher proportion of male patients was found in the B-UCSF group (76.2% vs 83.3% vs 93.3%, respectively; \( p = 0.031 \)). Causes of cirrhosis were comparable between the groups. The most common cause of HCC and cirrhosis for the 3 groups was hepatitis C virus (67.8%, 61.1%, and 66.7%, respectively), followed by a history of alcohol abuse. Patients in the B-UCSF group had more advanced underlying liver disease, with a median calculated MELD of 14 vs 12 (within Milan criteria) and 11 (W-UCSF) (\( p = 0.023 \)). Although the last pre-LT AFP levels were comparable between groups, significant differences were found when the peak AFP levels before transplantation were compared; 15.2, 10.2, and 57.2 ng/dL, respectively, reporting the highest levels in B-UCSF group (\( p < 0.0001 \)). Patients in the B-UCSF

Table 2. Univariate Analysis Between Within Milan Criteria, Within University of California San Francisco Criteria, and Beyond University of California San Francisco Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Within MC (( n = 239 ))</th>
<th>Within UCSF criteria (( n = 18 ))</th>
<th>Beyond UCSF criteria (( n = 45 ))</th>
<th>( p ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex, n (%)</td>
<td>182 (76.2)</td>
<td>15 (83.3)</td>
<td>42 (93.3)</td>
<td>0.031*</td>
</tr>
<tr>
<td>Caucasian, n (%)</td>
<td>198 (82.8)</td>
<td>14 (77.8)</td>
<td>35 (77.8)</td>
<td>0.860</td>
</tr>
<tr>
<td>Age, y, mean (SD)</td>
<td>58.3 (7.3)</td>
<td>56.4 (5.5)</td>
<td>59.1 (7.0)</td>
<td>0.427</td>
</tr>
<tr>
<td>Cirrhosis, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatitis C virus</td>
<td>162 (67.8)</td>
<td>11 (61.1)</td>
<td>30 (66.7)</td>
<td>0.841</td>
</tr>
<tr>
<td>Alcoholic</td>
<td>33 (13.8)</td>
<td>3 (16.7)</td>
<td>9 (20.0)</td>
<td>0.551</td>
</tr>
<tr>
<td>Nonalcoholic steatohepatitis</td>
<td>20 (8.4)</td>
<td>3 (16.7)</td>
<td>4 (8.9)</td>
<td>0.493</td>
</tr>
<tr>
<td>Cholestatic</td>
<td>5 (2.1)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Calculated Model for End-stage Liver Disease, median (range)</td>
<td>12 (6–41)</td>
<td>11 (6–22)</td>
<td>14 (6–49)</td>
<td>0.023*</td>
</tr>
<tr>
<td>Last AFP before LT, ng/mL, median (IQR)</td>
<td>9.8 (5–9,458)</td>
<td>5.5 (5–140)</td>
<td>10.6 (5–1,909)</td>
<td>0.778</td>
</tr>
<tr>
<td>AFP peak level, ng/mL, median (IQR)</td>
<td>15.2 (6.5–9,729)</td>
<td>10.2 (5.5–1,999)</td>
<td>57.2 (10.7–39,660)</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Child-Turcotte-Pugh class, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>120 (50.2)</td>
<td>11 (61.1)</td>
<td>23 (51.1)</td>
<td>0.762</td>
</tr>
<tr>
<td>B</td>
<td>82 (34.3)</td>
<td>4 (22.2)</td>
<td>13 (28.9)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>37 (15.5)</td>
<td>3 (16.7)</td>
<td>9 (20.0)</td>
<td></td>
</tr>
<tr>
<td>Time on waiting list, median (IQR)</td>
<td>100 (1–2,749)</td>
<td>92 (12–345)</td>
<td>113 (2–2,934)</td>
<td>0.433</td>
</tr>
<tr>
<td>Last locoregional therapy to LT interval, median (IQR)</td>
<td>90 (3–597)</td>
<td>120 (22–303)</td>
<td>135 (11–610)</td>
<td>0.0003*</td>
</tr>
<tr>
<td>Tumors, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>180 (75.3)</td>
<td>6 (33.3)</td>
<td>13 (28.9)</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>2</td>
<td>37 (15.5)</td>
<td>9 (50.0)</td>
<td>13 (28.9)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>22 (9.2)</td>
<td>3 (16.7)</td>
<td>5 (11.1)</td>
<td></td>
</tr>
<tr>
<td>&gt;3</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>14 (31.1)</td>
<td></td>
</tr>
<tr>
<td>No. of tumors, mean (SD)</td>
<td>1.3 (0.6)</td>
<td>1.8 (0.7)</td>
<td>2.9 (2.3)</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Largest tumor diameter, mean (SD)</td>
<td>2.7 (0.9)</td>
<td>4.5 (1.2)</td>
<td>5.1 (2.5)</td>
<td>&lt;0.0001*</td>
</tr>
</tbody>
</table>

*Significant.

AFP, \( \alpha \)-fetoprotein; LT, liver transplantation; MC, Milan criteria; UCSF, University of California San Francisco.
group also waited the longest after their last LRT (135 days) compared with within Milan criteria (90 days) and W-UCSF (120) groups (p = 0.0003). Wait time after listing was comparable among the groups. Patients in the within Milan criteria group had solitary tumors more frequently (75.3%) compared with W-UCSF (33.3%) and B-UCSF (28.9%) groups. As expected, the largest tumor diameter was also smaller in within Milan criteria group (2.7 cm) compared with the other 2 groups (4.5 and 5.1 cm, respectively) (p < 0.0001).

In the B-UCSF group, 31.1% patients had >3 tumors. On average, these patients had 2.9 (±2.3) tumors with mean ± SD largest diameter of 5.1 ± 2.5 cm. Nineteen (42.2%) patients in this group started with a total tumor burden of >8 cm. Of these 19 patients, 5 had a single tumor >8 cm, 7 had 2 to 3 tumors, and 7 had >3 tumors.

**Overall survival**

Thirty day, 1-, 3-, 5-, and 10-year overall survival rates in the full cohort of 302 transplanted HCC cases were 97.7%, 92.7%, 82.3%, and 64.0%, respectively. Median follow-up was 51.4 months. Overall 1-, 3-, 5-, and 10-year survival rates in the B-UCSF group (88.9%, 73.2%, 66.2%, and 49.7%, respectively) were comparable with those in the W-UCSF and within Milan criteria groups (94.4%, 94.4%, 85.8%, and 62.6%; and 93.2%, 83.0%, 74.1%, and 61.9%, respectively) (p = 0.29) (Fig. 2).

**Disease-specific survival**

Thirty day, 1-, 3-, 5-, and 10-year disease-specific survival rates in the full cohort of 302 transplanted HCC cases were 100.0%, 99.3%, 94.8%, 92.6%, and 90.6%, respectively. One-, 3-, 5-, and 10-year disease-specific survival rates in the B-UCSF group (97.7%, 92.3%, 92.3%, and 92.3%, respectively) were also comparable with those in W-UCSF and within Milan criteria groups (100%, 100%, 90.9%, and 90.9%; and 99.6%, 94.8%, 92.9%, and 90.5%, respectively) (p = 0.95) (Fig. 3).

**Deaths and hepatocellular carcinoma-related deaths**

Ninety-two (30.5%) patients died during the observation period. In the B-UCSF group, 16 (35.6%) patients have died. Causes of death were recurrent HCC (n = 3) at 0.6, 1.1, and 2.2 years from transplantation, cancer other than HCC (n = 2), multisystem failure and sepsis (n = 4), cardiac (n = 1), gastrointestinal bleeding (n = 1), respiratory failure (n = 1), and miscellaneous (n = 4).

In the W-UCSF group, 4 (22.2%) patients have died. Causes of death were recurrent HCC (n = 1) at 3.9 years from the transplantation, liver graft failure (n = 2), and cardiac (n = 1).

In the within Milan criteria group, 72 (30.1%) patients have died. Causes of death were recurrent HCC (n = 16) at a median interval of 1.9 (range 1.0 to 6.4) years from transplantation, intraoperative bleeding (n = 2), cancer other than HCC (n = 6), multisystem failure and
sepsis (n = 6), cardiac (n = 5), gastrointestinal bleeding (n = 2), respiratory failure (n = 5), liver graft failure (n = 8), renal failure (n = 6), hepatorenal failure (n = 3), and miscellaneous (n = 13).

**Disease-free survival**
Thirty day, 1-, 3-, 5-, and 10-year disease-free survival rates in the full cohort of 302 transplanted HCC cases were 97.7%, 89.0%, 80.3%, 70.6%, and 58.1%, respectively. One-, 3-, 5-, and 10-year disease-free survival rates in the B-UCSF group (84.4%, 73.5%, 62.8%, and 44.9%, respectively) were also comparable with those in W-UCSF and within Milan criteria groups (94.4%, 87.2%, 87.2%, and 63.6%; and 89.5%, 80.9%, 70.7%, and 59.6%, respectively) (p = 0.28) (Fig. 4).

**Hepatocellular carcinoma recurrence**
Thirty day, 1-, 3-, 5-, and 10-year disease recurrence rates in the full cohort of 302 transplanted HCC cases were 0.0%, 4.5%, 7.3%, 10.9%, and 11.6%, respectively. One-, 3-, 5-, and 10-year HCC recurrence rates in the B-UCSF group (6.9%, 6.9%, 12.1%, and 12.1%, respectively) were also comparable with those in W-UCSF and within Milan criteria groups (0.0%, 7.7%, 7.7%, and 7.7%; and 4.4%, 7.2%, 10.9%, and 11.9%, respectively) (p = 0.85) (Fig. 5).

**Disease recurrence**
Hepatocellular carcinoma developed in 27 patients. Four (8.9%) were B-UCSF patients, 1 (5.6%) was a W-UCSF patient, and 22 (9.2%) were within Milan criteria patients.

In the 4 B-UCSF patients with disease recurrence, 2 recurrences were in the lungs, 0.6 and 2.2 years after the transplantation. One patient had more widespread recurrent disease in the lumbar vertebrae, lungs, and bilateral 12th ribs, 0.5 years after the transplant. Hepatocellular carcinoma developed in the transplanted liver of the other patient after 4.2 years.

Recurrent disease in the soft tissue over the left scapular region developed in 1 W-UCSF patient 3 years after transplantation. Of the 22 within Milan criteria patients with recurrent disease, recurrence was restricted to the transplanted liver in 4 cases. These recurrences were observed 0.5, 1.1, 4.0, and 5.4 years after transplantation. Of the remaining 18, 9 recurrences occurred during the first year after the transplantation (4 in lymph nodes, 2 in bone, 1 in lung + liver, 1 in lung + both adrenals, and 1 with widespread disease).

Five recurrences occurred after the first year, but before completing 3 years post-orthotopic liver transplantation (2 liver + vertebrae, 1 omentum, 1 lymph nodes, 1 widespread). The last 4 recurrences occurred between 3 and 5.4 years post-orthotopic liver transplantation (2 in lungs, 1 on the chest wall, and 1 in the mediastinum).

**Graft survival**
Thirty-day, 1-, 3-, 5-, and 10-year graft survival rates in the full cohort of 302 transplanted HCC patients were 96.0%, 90.3%, 80.0%, 72.2%, and 59.6%, respectively. One-, 3-, 5-, and 10-year graft survival rates in the B-UCSF group (86.7%, 73.7%, 66.7%, and 50.0%, respectively) were also comparable with those in W-UCSF and within Milan criteria groups (94.4%, 87.2%, 87.2%, and 63.6%; and 89.5%, 80.9%, 70.7%, and 59.6%, respectively) (p = 0.28) (Fig. 4).
criteria groups (94.4%, 94.4%, 85.9%, and 62.6%; and 90.7%, 80.0%, 72.2%, and 60.9%, respectively) (p = 0.39) (Fig. 6).

α-Fetoprotein levels
Historical AFP levels before transplantation were analyzed in our cohort. Sixteen (5.3%) patients presented with AFP levels >1,000 ng/mL at some point during the evaluation for LT. Seven of 239 (2.9%) of these were in the within Milan criteria group, 1 of 18 (5.6%) was from the W-UCSF group, and 8 of 45 (17.8%) patients were from the B-UCSF group. Nine of these 16 LTs occurred before 2009. All 16 received LRT between this peak AFP report and LT.

After LRT, AFP level came down to <500 ng/mL before LT in 6 patients (2 from the within Milan criteria group, 1 W-UCSF case, and 3 from the B-UCSF group). In 1 of these patients (from within Milan criteria group), metastatic disease developed in the lymph nodes within 5 months post-LT; the patient received radiation therapy and had post-LT survival of 4 years. Median overall survival of these 6 patients was 5.6 years.

Four patients from the B-UCSF group had partial response as their AFP level came down to <1,000 ng/mL, but could not be brought down to <500 ng/mL. There was no recurrence in these 4 patients. Their median post-LT survival is 5.1 years.

Despite LRT, 6 patients’ AFP levels remained >1,000 ng/mL. One of these was in the B-UCSF group and the other 5 were in within Milan criteria group. Five of the 6 cases were performed before 2009. Recurrent HCC developed in 1 of these patients (within Milan criteria case) 10 months after transplantation in his transplanted liver as well as in the left lung and he died 20 months later, despite receiving cryoablation of the liver lesion and resection of the lung lesions. Median overall survival of these patients was 8.1 years.

Of the 286 patients who never had their AFP level >1,000 ng/mL, 10 had their last pre-LT AFP >500 ng/mL. Four of them were in B-UCSF group and 6 were in within Milan criteria group. No LRT was given in the interval between this AFP report and LT. Recurrence developed in 2 of these patients from the within Milan criteria group. One of the 2 patients showed recurrent disease in bone 4 months post-LT, was advised chemoradiation but chose not to pursue additional treatment and died with post-LT survival of 1 year. The other patient’s recurrent disease was restricted to the liver, discovered 14 months post-LT. He received RFA and is alive and disease-free 12+ years after LT.

Cox regression analysis
As an independent factor, beyond–Milan criteria disease was not found to significantly affect survival (overall, disease-specific, disease-free, and graft survival) or disease recurrence. There were no differences in these outcomes between B-UCSF and W-UCSF groups. The variables that significantly affected overall survival were hepatitis C virus cirrhosis (p = 0.031; hazard ratio = 1.689; 95% CI 1.049 to 2.725) and alcohol-associated cirrhosis (p = 0.018; hazard ratio = 1.869; 95% CI 1.114 to 3.135). Longer interval between last LRT and transplantation (>3 months) was independently associated with improved disease-free survival (p = 0.038; hazard ratio = 1.529; 95% CI 1.022 to 2.283). Waiting on the list for longer than 6 months showed a trend to affect the time to disease recurrence, but with a p value of 0.054, failed to reach the level of significance.

DISCUSSION
Treatment of patients with cirrhosis and HCC within Milan criteria is well established in Western countries. After the introduction of these criteria by Mazzaferro and colleagues11 in 1996, liver transplantation has become the standard treatment for patients with early-stage HCC. Patients within Milan criteria receive MELD priority points, allowing for prioritization for transplantation, usually within 6 to 12 months. However, patients who exceed these criteria do not receive automatic priority on the waiting list, which can lead to dropout and death from progression of disease. The application of LRT to achieve reduction of the tumor stage allows patients...
outside these criteria to potentially access transplantation. In the current report, we demonstrate that patients beyond Milan criteria can be successfully downstaged to within Milan criteria, with post-transplantation survival results that mirror those for patients initially presenting within Milan criteria. In addition, a priori restriction of candidates for possible downstaging was not required, as long as HCC was confined to the liver, and this included patients with AFP >1,000 ng/mL and patients with evidence of intrahepatic portal and hepatic venous tumor thrombus.

In the current report, we analyzed a total of 210 patients with HCC beyond Milan criteria who received LRT with the goal of downstaging to within Milan criteria; 89 (42%) of these patients underwent successful downstaging and 63 (30%) subsequently underwent LT. We performed an analysis of factors that could have predisposed to a greater probability of downstaging, comparing these 89 patients with the 121 patients that were not successfully downstaged. Based on our results, younger patients appear to have a greater probability of downstaging. Interestingly, meeting UCSF criteria at initial presentation was not a limiting factor to achieve downstaging in our study and did not impact post-transplantation survival results. In a related study investigating the effect of LRT on long-term outcomes in 501 patients (both within and outside of Milan criteria), Agopian and colleagues analyzed the response to LRT on outcomes after LT. Patients achieving complete response with LRT had post-transplantation recurrence risk of 2% compared with 18% for those who did not achieve complete response. They found that tumor size was not predictive of achieving complete response after LRT, and increasing tumor number did lower the likelihood of achieving complete response. In our series, tumor size did not have a significant effect on downstaging success, although patients who did not achieve downstaging tended to have larger tumor sizes compared with those who were successfully downstaged.

We did not find differences when we compared the number of LRTs received in each group based on the likelihood of achieving downstaging. Of note, we did not follow a single protocol for downstaging, but instead, each case was individually evaluated by a multidisciplinary team and the type of LRT and number of sessions was decided on a case-by-case basis. Despite the desirability of a single standardized approach for LRT, no method has yet been proven superior based on studies to date, and our limited use of Y-90 and RFA do not allow us to comment on this aspect of LRT. Agopian and colleagues described lower HCC recurrence rates after obtaining complete response with direct ablation therapies (eg RFA), although they found that the combination therapy with TACE and ablation was the strongest treatment-specific factor in predicting complete response. In our study, the most commonly used LRT was TACE (81.4%), 23 (11%) patients received Y-90 as unique LRT, and 16 patients (7.6%) received combination therapy with TACE + Y-90. Ten patients receiving TACE and 1 patient receiving Y-90 also received RFA as a combination therapy. There is a high variability of LRT techniques, and so far there is insufficient evidence to suggest superiority of any single or combination technique.

Patients with HCC within Milan criteria at initial presentation who underwent LT were used as our control group to compare the outcomes obtained after transplantation in patients W-UCSF and B-UCSF who underwent downstaging LRT. All groups represented a homogeneous distribution of age, race, cause of cirrhosis, MELD score, and Child-Turcotte-Pugh class. As expected, patients in the B-UCSF group initially presented with a greater number of tumors and larger tumor size than other groups (within Milan criteria and B-UCSF). We found that the time between LRT and LT was greater in the B-UCSF group. Several studies have described a minimum observation period of 3 months after downstaging and before LT as a tool to monitor tumor biology and to select tumors with more favorable prognosis. We found that the time interval between the last LRT and transplantation (≤3 vs >3 months) independently affected disease-free survival and patients with longer waiting times had greater disease-free survival (p = 0.04). In our practice, an imaging study (CT/MRI) was usually performed 4 to 6 weeks after LRT and, if downstaging was achieved, patients were able to undergo listing for LT. With the recently enacted United Network for Organ Sharing policy of a mandatory 6-month delay between listing for LT for HCC and granting of priority points, this will result in a more than 6-month delay in all patients between last LRT and transplantation for downstaged and within Milan criteria patients.

The AFP levels have been suggested to play an important role when selecting patients as candidates for downstaging and as a predictor of treatment response. Most studies that have compared patients that achieved downstaging with patients that were unsuccessful, showed higher levels of AFP in the nondownstaged group. Barakat and colleagues found that the mean AFP level at the time of presentation was significantly higher in the nondownstaged group. On the other hand, De Luna and colleagues found no significant predictors for successful downstaging to Milan criteria, including AFP levels. Our data demonstrate higher levels of AFP in the nondownstaged group and AFP levels before the first LRT showed a trend to be a predictor of downstaging failure, but did not reach the level of
significance (45.7 vs 16.7 ng/mL; p = 0.06). Yao and colleagues reported that an AFP level >1,000 ng/mL was a predictor of treatment failure in the UCSF study, and an AFP level >1,000 ng/mL was included as an additional exclusion criterion (unless the AFP level decreased to <500 ng/mL with LRT) at the US national conference. Other groups have suggested that an AFP level >400 ng/mL should be used as an exclusion criterion for LT. Although the best AFP cutoff in predicting prognosis after LT is still unclear, an AFP >1,000 ng/mL has been associated with worse outcomes after LT for HCC within Milan criteria. We did not use AFP levels as exclusion criteria in the current study. Other reports have compared patients beyond Milan criteria and patients within Milan criteria and describe significantly higher AFP levels at the time of diagnosis and pretransplantation in patients beyond Milan criteria. When the last AFP levels pretransplantation were analyzed in our cohort, we did not find significant differences among the groups (within Milan criteria, W-UCSF, and B-UCSF), although significant differences were found when the AFP peak levels were compared, with higher levels in the B-UCSF group (p < 0.0001). This can be explained by a decrease in AFP levels after successful downstaging and before LT in this group. A higher proportion of patients in the B-UCSF group had AFP levels >1,000 ng/mL at least once compared with those in W-UCSF and within Milan criteria groups; 17.9% vs 5.6% vs 2.9%, respectively. Despite these findings, none of the patients who presented with AFP levels >1,000 ng/mL in the B-UCSF group had recurrence of disease. Of the 6 patients (1 within Milan criteria, 5 B-UCSF) whose last AFP was >1,000 ng/mL before transplantation, recurrent HCC developed in 1, and this patient was in the within Milan criteria category. Of the 6 patients (2 within Milan criteria, 1 W-UCSF, and 3 B-UCSF) who initially had AFP >1,000 ng/mL but eventually decreased to <500 ng/mL, there was 1 patient with recurrence in the within Milan criteria category. Of the patients with initial AFP level >1,000 ng/mL, but subsequently had AFP between 500 and 1,000 ng/mL (who would be excluded from LT under the current United Network for Organ Sharing proposal), 4 (all B-UCSF) received transplants with no recurrence observed to date. We question the United Network for Organ Sharing proposal that allows for LT exception points for AFP up to 1,000 ng/mL, but requires those >1,000 ng/mL to fall (and remain) <500 ng/mL, based on lack of sufficient evidence to exclude this group from consideration of transplantation.

There are several limitations to the current study, including the lack of a uniform protocol for LRT in use for all patients with HCC. However, all patients underwent a similar prospective multidisciplinary team review to determine the LRT approach and this was true for patients within Milan criteria and beyond Milan criteria. In addition, like all retrospective studies, there exists a possibility of unmeasured selection bias that might have affected patient selection for LT; however, data were obtained from our prospective transplant database and patient follow-up after LT is thorough and stringent, making determinations of HCC recurrence highly accurate.

Several proposals have arisen in an attempt to extend the Milan criteria to offer the option of LT to a greater number of patients with advanced HCC. In 2001, the UCSF group first proposed the UCSF criteria and showed results comparable with those obtained in patients with stage II HCC. Since that time, other centers have evaluated the extension of these criteria and have evaluated results in patients with higher tumor burden. Our study compares survival obtained in patients beyond Milan criteria but W-UCSF criteria, as well as patients B-UCSF criteria who underwent LRT until successful downstaging was achieved before LT. Both groups, W-UCSF and B-UCSF, demonstrated an overall survival comparable with the survival obtained in patients within Milan criteria and an approximately 10% HCC recurrence risk at 5 years of follow-up. After comparing graft, disease-free, and disease-specific survival, in addition to recurrence rates, our results support the concept that regardless of tumor size and number of tumors at presentation, HCC patients can undergo LRT as an attempt of downstaging. Once patients show downstaging to within Milan criteria based on cross-sectional imaging, patients who initially exceed Milan criteria can undergo LT with outcomes comparable with those obtained in stage II patients. Based on these results, our belief is that there should be no limits for the access of these patients to LRT a priori and, regardless of the number and size of tumors, attempting downstaging with LRT with a goal of achieving Milan criteria status and listing for transplantation is reasonable.

CONCLUSIONS
Patients with beyond—Milan criteria HCC who are otherwise candidates for LT should undergo aggressive attempts at downstaging, without a priori exclusion. This highly selective approach allows for excellent long-term results, similar to patients presenting with earlier-stage disease.

Author Contributions
Study conception and design: Chapman, Garcia-Aroz, Vachharajani, Fowler, Saad, Lin, Wellen, Tan, Khan, Doyle
Acquisition of data: Chapman, Garcia-Aroz, Vachharajani, Fowler, Saad, Lin, Wellen, Tan, Khan, Doyle
Analysis and interpretation of data: Chapman, Garcia-Aroz, Vachharajani, Fowler, Saad, Lin, Wellen, Tan, Khan, Doyle
Drafting of manuscript: Chapman, Garcia-Aroz, Vachharajani, Fowler, Saad, Lin, Wellen, Tan, Khan, Doyle
Critical revision: Chapman, Garcia-Aroz, Vachharajani, Fowler, Saad, Lin, Wellen, Tan, Khan, Doyle
REFERENCES