ORIGINAL ARTICLE

Treatment of hepatocellular carcinoma beyond the Milan criteria. A weighted comparative study of surgical resection versus chemoembolization

Simone Famularo[1,2], Stefano Di Sandro[3], Alessandro Giani[1,2], Davide P. Bernasconi[1,4], Andrea Lauterio[3], Cristina Ciulli[1,2], Antonio G. Rampoldi[5], Rocco Corso[6], Riccardo De Carlis[3,7], Fabrizio Romano[1,2,8], Marco Braga[1,2,8], Luca Gianotti[1,2,8] & Luciano De Carlis[1,3,8]

[1]School of Medicine and Surgery, University of Milano - Bicocca, [2]Department of Surgery, ASST - San Gerardo Hospital, Monza, [3]Department of General Surgery and Transplantation – ASST - Grande Ospedale Metropolitano Niguarda, Milan, [4]Center of Biostatistics for Clinical Epidemiology, School of Medicine and Surgery, University of Milano - Bicocca, Monza, [5]Department of Radiology, ASST - Grande Ospedale Metropolitano Niguarda, Milan, [6]Department of Radiology, ASST - San Gerardo Hospital, Monza, [7]Department of Surgical Sciences, University of Pavia, Pavia, and [8]International Center for Digestive Health, University of Milan-Bicocca, Milan, Italy

Abstract

Background: Optimal treatment of hepatocellular carcinoma (HCC) beyond the Milan criteria (MC) is debated. The aim of the study was to assess overall-survival (OS) and disease-free-survival (DFS) for HCC beyond MC when treated by trans-arterial-chemoembolization (TACE) or surgical resection (SR).

Method: between 2005 and 2015, all patients with a first diagnosis of HCC beyond MC(1 nodule>5 cm, or 3 nodules>3 cm without macrovascular invasion) were evaluated. Analyses were carried out through Kaplan–Meier, Cox models and the inverse probability weighting (IPW) method to reduce allocation bias. Sub-analyses have been performed for multinodular and single large tumors compared with a MC-IN cohort.

Results: 226 consecutive patients were evaluated: 118 in SR group and 108 in TACE group. After IPW, the two pseudo-populations were comparable for tumor burden and liver function. In the SR group, 1–5 years OS rates were 72.3% and 35% respectively and 92.7% and 39.3% for TACE ($p = 0.500$). The median DFS was 8 months (95%CI:8–9) for TACE, and 11 months (95%CI:9–12) for SR ($p < 0.001$). TACE was an independent predictor for recurrence (HR 1.5; 95%CI: 1.1–2.1; $p = 0.015$). Solitary tumors > 5 cm and multinodular disease had comparable OS and DFS as Milan-IN group ($p > 0.05$).

Conclusion: Surgery allowed a better control than TACE in patient bearing HCC beyond MC. This translated into a significant benefit in terms of DFS but not OS.

Received 9 June 2019; accepted 16 December 2019

Correspondence

Luca Gianotti, Department of Surgery (4° piano A), San Gerardo Hospital, Via Pergolesi 33, 20900, Monza, Italy. E-mail: luca.gianotti@unimib.it

Introduction

Patients bearing hepatocellular carcinoma (HCC) and exceeding the Milan criteria (MC)[1] (single tumor < 5 cm or multinodular disease up to 3 lesions, each < 3 cm in diameter) have a poor long-term prognosis with a median survival of nearly 2 years and a 5-year survival rate of 14–25%.[2–4] Given for granted that those patients received the best available treatment, this statistic is significantly worse if compared with that of patients who are fitting the MC.[5,6] According to the 2018 guidelines for the management of HCC by the European Association for the Study of the Liver (EASL),[7] MC-out patients, in the absence of macrovascular invasion, belong prevalently to the intermediate stage as per the BCLC staging system. The recommended gold-standard management for this class was established to be the trans-arterial chemoembolization (TACE) of the tumor. However, there is a not negligible body of evidence[8,9] suggesting that

This study was presented at the HCC Summit 2019 (Lisbon, 14–16 February 2019) and has been awarded the Young Investigator Award (full bursary). Another oral presentation has been given during the 13th congress of the European-African Hepato-pancreato-biliary association (E-AHPBA) meeting 2019 (Amsterdam, 2–5 June 2019).

in selected patients, surgical resection (SR) can offer significant survival benefits when compared with palliative treatments as TACE. While within the MC the surgical approach is well accepted, the real impact of curative rather than palliative therapies beyond MC are still debated.The primary aim of our study was to compare overall survival (OS) of patients treated with SR or TACE for HCC beyond MC. Disease-free survival (DFS) was the secondary endpoint. Given the predictable difference between the two groups we adopted the inverse probability weighting approach to minimize the potential selection bias in analysing the association between treatments and outcomes. Furthermore, to better understand the impact of surgery, the surgical MC-Out cohort was compared with a cohort of MC-In patients, in which the operation represents the standard of care.

Methods

Study overview and treatments

This retrospective study evaluated patient data collected prospectively in two Italian centers (ASST - San Gerardo Hospital, Monza, and ASST - Grande Ospedale Metropolitano Niguarda, Milan) and anonymized prior to the retrospective analysis. The study protocol followed the ethical guidelines of the 1975 Declaration of Helsinki (as revised in Brazil 2013). The local Ethical Committee review of the protocol deemed that formal approval was not required owing to the retrospective, observational and anonymous nature of this study. Results are reported according to principles of Strengthening the Reporting of Observational Studies in Epidemiology (STROBE).[11] Data collection was performed using an electronic database system in both centers. The submitted data were then checked centrally, at San Gerardo Hospital and, when missing data were identified, the local investigator was contacted and asked to complete the records. Once examined, the record was accepted into the dataset for analysis.

Details on the surgical technique have been previously described.[12,13] TACE was performed as follows: a vascular catheter was inserted through the femoral artery using the Seldinger technique to the hepatic artery, and hepatic angiography was carried out. Hepatic arteriography was performed to identify the feeding artery of the liver tumor. The catheter, usually a 4F or 5F RH catheter (either Simmons [Cook, Bloomington, Ind] or Cobra [Terumo Medical, Somerset, NJ] catheters were used) was inserted into the tumor feeding artery as close as possible to the tumor. Microcatheters (Progreat, Terumo Co, Kanagawa, Japan) were used to catheterize in a superselective way the feeding artery if needed. Once the operator selected the final catheter position for TACE, intra-arterial chemotherapy was performed by injection of 10–12 mL of iodized oil (Lipiodol Ultra Fluide; Laboratoire Guerbet, Roissy, France) mixed with an emulsion of 50 mg of doxorubicin hydrochloride into the feeding artery. Embolization was performed by means of a mixture of iobitridol (Xenetix 350; Guerbet, Aulnay, France) and 1-mm-diameter absorbable gelatin sponge particles (Spongostan; Ferrosan, Søborg, Denmark). The regimen of chemoembolization was adjusted according to liver function and peripheral leukocyte or platelet levels. Gelatin was administered afterward for additional embolization. The radiologic response was evaluated 4–6 weeks thereafter by contrast-enhanced multidetector computed tomography or magnetic resonance.

Patient selection and study design

All consecutive adult patients (age $\geq$ 18 years) with histologically or radiologically proven HCC who underwent SR or TACE at the two institutions from January 2005 to December 2015 were evaluated for this study. Inclusion criteria were: (i) first diagnosis of HCC without any previous disease treatment; (ii) Eastern Cooperative Oncology Group (ECOG) performance status 0 or 1. Exclusion criteria were (i) macrovascular tumor invasion; (ii) Child-Pugh class C; (iii) patients on a waiting list for transplantation or receiving any bridge treatment to transplant. Patients were then divided according to MC criteria (IN or OUT). Finally, patients were analysed according to the treatment allocation: SR or TACE.

The indication to surgery or TACE was assessed during multidisciplinary meetings involving surgeons, hepatologists, oncologists, radiologists, interventional radiologists, infectiologists. Patient-tailored treatment allocation was based on patient comorbidities, previous abdominal operations, underlying liver damage and presence of bilobar disease.

The primary analysis encompassed MC-Out patients who underwent surgery or TACE.

In a secondary analysis, we compared MC-out patients with a reference group (MC-in) in which surgery is the recognized standard of care. This was done to better understand whether surgery may be considered an acceptable treatment option also in MC-out. We further stratify this cohort in single large tumors (SR-SN) and multinodular tumors (SR-MN), and we compared these subgroups with MC-IN patients with similar characteristics. The decision flow chart summarizing the different group comparisons is depicted in Fig. 1.

Outcomes

The primary endpoint was to compare the overall-survival (OS) in patients undergoing surgery or TACE in a MC-OUT cohort. The secondary endpoint was disease-free-survival (DFS). Risk factors for OS and DFS were also evaluated.

As a subgroup analysis, we also evaluated the survival outcomes of surgery and TACE in patients stratified as per the Bolondi *et al.* classification[10]: B1 (Child 5–7; beyond MC but within Up-to-7; ECOG 0; no portal vein thrombus); B2 (Child 5–6; beyond MC and Up-to-7; ECOG 0; no portal vein thrombus); B3 (Child 7; beyond MC and Up-to-7; ECOG 0; no portal vein thrombus) and B4 (Child 8–9; beyond MC and beyond or within Up-to-7; ECOG 0–1; no portal vein thrombus).

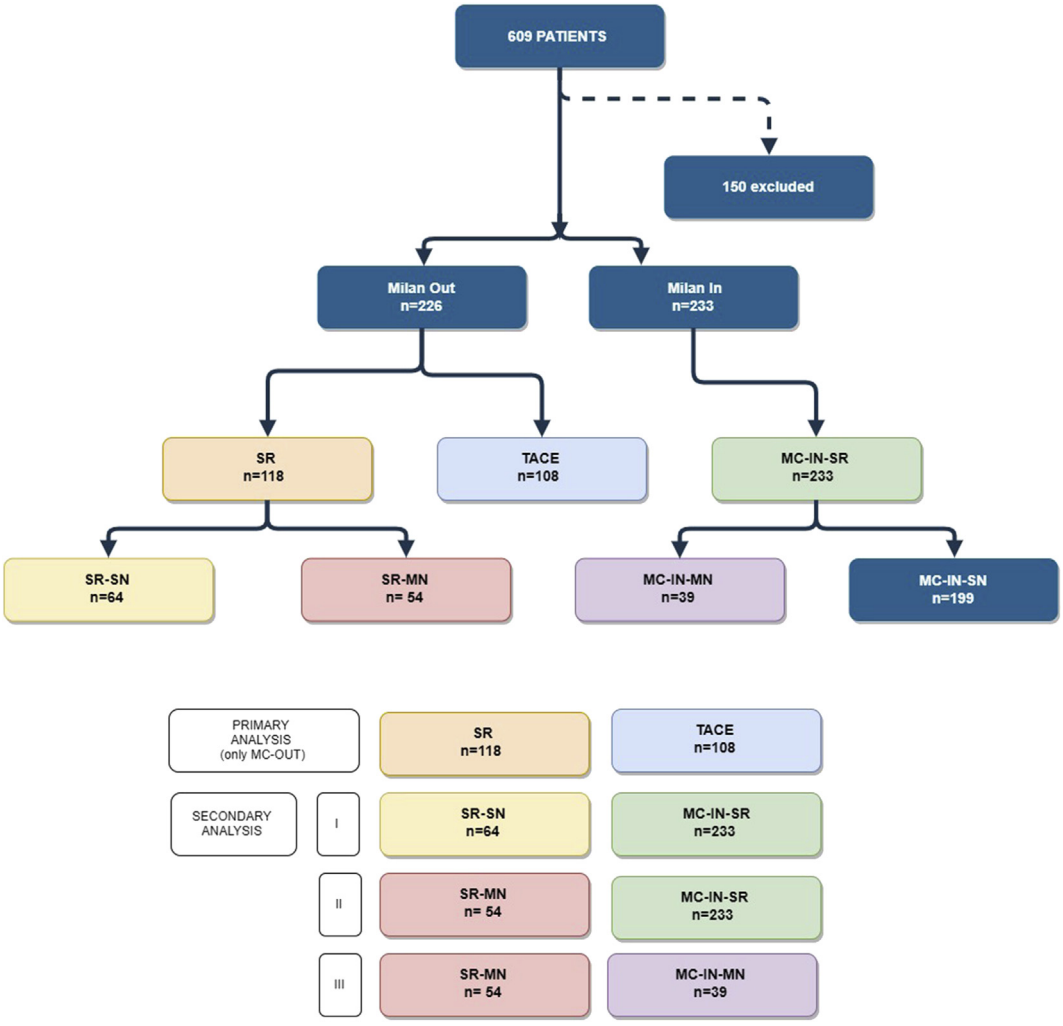

Figure 1 Flowchart of allocation and subgroup comparisons. SR Milan Out Surgical Resection Group; TACE Milan Out trans-arterial chemoembolization; MC-IN-SR Milan In Surgical Resection Group; MC-OUT Milan Criteria – OUT; SR-SN Milan Out Surgical Resection Single Large Nodule group; SR-MN Milan Out surgical resection multinodular group; MC-IN-MN Milan In multinodular group; MC-IN-SN Milan In Single Large Nodule group

Variables

Age, sex, Eastern Cooperative Oncology Group (ECOG) performance status[14] and liver function at presentation were recorded and evaluated at the first visit. In particular, the presence of cirrhosis and its severity was evaluated by expert hepatologists during the disease work-up. BCLC grade was estimated after radiological evaluation. Model for end-stage liver disease (MELD) score, Child-Pugh score, were calculated on the basis of preoperative serum biochemical values and clinical examination. Serum alpha-fetoprotein (α-FP) was also part of preoperative evaluation. The number and diameter of nodules were assessed through preoperative radiologic imaging and confirmed by intraoperative ultrasound.

Follow-up

All patients were followed using a local protocol including measurement of serum α-fetoprotein, abdominal ultrasound, contrast computed tomography (CT) or magnetic resonance imaging (MRI), and office visits as previously described.[15] Briefly, each patient was followed up every 3 months for the first two years and then every six months. OS was defined as the time interval in months from surgery or first TACE to death; if alive, patient data were censored at the last visit available. DFS was defined as the time interval in months from surgery or TACE to recurrence or death. In case of no recurrence or death, data were censored at the date of the last available follow-up. Patient surveillance was closed at the end of March 2018.

Statistical analysis

Descriptive statistics are expressed as median and interquartile range (IQR) for continuous variables and as number and proportion for categorical variables. The distribution of baseline factors was compared between treatments (SR vs. TACE) using Mann–Whitney or Fisher test as appropriate. The probability of OS over time was estimated using the Kaplan–Meier method. Stratified curves and the log-rank test were used to evaluate the association of prognostic factors with OS. Moreover, uni- and multivariate Cox regression models were fitted to the data on factual population. The inverse probability weighting (IPW) approach was applied to minimize potential selection bias in analysing the association between treatment and OS, as follows: logistic regression was carried out to estimate the probability of receiving SR of TACE depending on patient and tumor characteristics. These were chosen among factors associated to both treatment and OS at the univariate analysis or based on clinical knowledge. Subsequently, the outcome of each patient was weighted by the inverse of the probability of the treatment actually received creating a pseudo-population of doubled size that resembles the counterfactual situation were all patients receive both treatments. In the weighted sample, measured confounders should be balanced between treatment groups. As a consequence, the standardized difference between groups of all variables included in the predictive model for treatment was negligible on the weighted sample compared to the standardized difference on the original sample. Finally, treatment-specific Kaplan–Meier curves of OS were estimated on the weighted sample data. Analogous analyses were performed on the DFS end-point. All statistical tests were two tailed and a 5% significance level was considered. All the analysis was carried out using R software version 3.5.1.

Results

From 2005 to 2015, 609 patients were evaluated during multidisciplinary meetings. Of them, 226 patients were outside the MC, while 233 were classified as within MC. The other 150 were excluded according to the criteria of the study. Within the cohort of MC-out patients, 118 received SR and 108 TACE. All patients in the MC-IN group underwent surgical resection.

SR and TACE in the MC-out cohort were different for several variables at univariate analysis: surgical group was more frequently female (26.3% vs 14.8%, p:0.048) and with a minor incidence of diabetes (20.3% vs 35.4%, p: 0.023). Moreover, the TACE group had an higher rate of cirrhotic patients (90.7% vs. 74.6% in surgery group, $p = 0.002$), a higher median MELD score (9 [IQR 8–11] versus 8 [6–10], $p < 0.001$), a median INR (1.17 [IQR 1.08–1.25] vs. 1.08 [IQR 1.00–1.18]), a higher frequency of HCV infections (61.1% vs 38.1%, p: 0.001) and also had a higher median number of nodules (3 [IQR 2–4] versus 1 [IQR 1–2] than in the surgery group, $p < 0.001$). Moreover, the TACE group had a higher rate of bilobar disease (46.7% vs 19.3%, $p < 0.001$). Conversely, the surgery group had an higher rate of tumors > 5 cm (72.9% vs 27.4% in the TACE group, $p < 0.001$). Data are summarized in Table 1.

To understand the impact of each significantly different variables on survival and treatment allocation, we performed a Cox univariate regression. Part of these variables, which may be associated to the tumor burden as well as estimators of worse underlying liver damage, where identified as factors with

Table 1 Univariate analysis comparing baseline characteristics of the two treatment groups

	Surgery (N = 118)	Tace (N = 108)	P
Age, years (median [IQR])	68.0 [56.0, 74.0]	68.0 [57.0, 77.0]	0.405
Female (%)	31 (26.3)	16 (14.8)	0.048
ECOG Performance status (%)			0.006
0	101 (85.5)	82 (75.9)	
1	11 (9.3)	0 (0.0)	
(missing)	(6)	(28)	
Diabetes (%)	24 (20.3)	29 (35.4)	0.023
Cardiopathy (%)	17 (14.4)	10 (12.2)	0.681
Pneumopathy (%)	11 (9.3)	0 (0.0)	0.003
Nephropathy (%)	3 (2.5)	0 (0.0)	0.271
Cirrhosis (%)	88 (74.6)	98 (90.7)	0.002
Child B (%)	20 (16.9)	20 (18.5)	0.862
MELD (median [IQR])	8.0 [6.0, 10.0]	9.0 [8.0, 11.0]	<0.001
HCV positive (%)	45 (38.1)	66 (61.1)	0.001
HBV positive (%)	30 (25.4)	17 (15.7)	0.1
Total bilirubin, mg/dL (median [IQR])	0.80 [0.50, 1.20]	0.90 [0.60, 1.50]	0.029
INR (median [IQR])	1.08 [1.00, 1.18]	1.17 [1.08, 1.25]	<0.001
Alpha-Feto protein (median [IQR])	27.9 [5.9, 415.7]	18.3 [8.1, 137.6]	0.51
Number of nodules (median [IQR])	1.00 [1.00, 2.00]	3.00 [2.00, 4.00]	<0.001
Tumor size, cm (median [IQR])	6.00 [5.00, 8.85]	4.00 [3.30, 5.50]	<0.001
Tumor size > 5 cm (%)	86 (72.9)	29 (27.4)	<0.001
Bilobar disease (%)	21 (19.3)	50 (46.7)	<0.001
Bolondi classification (%)			0.07
B1	48 (40.7)	50 (46.3)	
B2	56 (47.5)	43 (39.8)	
B3	10 (8.5)	4 (3.7)	
B4	4 (3.4)	11 (10.2)	

ECOG: Eastern Cooperative Oncology Group; MELD: model for end stage liver disease; HCV; hepatitis C virus; HPB: hepatitis B virus; INR: International Normalized Ratio.

increased risk of selection bias. The IPW method was used to weight the following confounders: presence of cirrhosis, MELD score, bilirubin, INR, number of nodules >1, tumor size > 5 cm. By this statistical approach we obtained two pseudo-populations (210.62 and 204.87 patients for SR and TACE respectively) with a reasonable certainty to compare the treatment effect on survival without being affected by liver or tumor related differences at the baseline, as measured by the reduction of the standard difference before and after weighting (Table 2). Supplementary Fig. 1(a and b) shows the distribution of the probabilities of TACE estimated by the logistic regression model and the distribution of the inverse probability of treatment weights, respectively, in the two factual treatment groups. The mean of the weights was close to ideal value 2 (1.84 and 2.15 in the surgery and TACE group, respectively).

Among the observed population, the median OS was 42 months (95%CI: 28–64) and 34 months (95%CI: 27–41) for SR and TACE respectively (p = 0.200). The median DFS for SR was 12 months (95%CI: 9–17) and 8 months (95%CI: 8–9) for TACE (p < 0.001). After IPW the results were similar. In the SR group, 1-3-5 year OS rates were 72.3%, 46.7% and 35% respectively and 92.7%, 59.4% and 39.3% for TACE (log-rank test = 0.500). In terms of DFS, 1-3-5 year survival rates were 44%, 19.8% and 15.3% for SR and 36.5%, 0.5% and 0.2% for TACE respectively (p < 0.001). Kaplan–Meier survival curves before and after IPW are depicted in Fig. 2.

Prognostic factors for OS and DFS

Table 3 shows the estimates of uni- and multivariate Cox models, fitted on factual population, to evaluate the impact of potential risk factors on mortality and recurrence. According to the results of the adjusted analysis, Child-B (HR 2.219; 95%CI: 1.379; 3.572; p = 0.001) and age (HR per one-year increase 1.172; 95%CI: 1.000; 1.373; p = 0.050) were significant predictors of OS. A poor liver function, (Child-B) (HR 1.822; 95%CI: 1.182; 2.808; p = 0.007), and TACE treatment (HR 1.535; 95%CI: 1.085; 2.171; p = 0.015) were significant predictors of worse DFS.

Surgery versus TACE after stratification according to the Bolondi classification

Survival curves stratified by the Bolondi *et al.* classification are depicted in Supplementary figure 2. B1 patients had a median OS of 32 months (95%CI: 27–43) after surgery and a median of 43 months (95%CI: 41–70) after TACE (log-rank test p = 0.70). The median OS for B2 was 39 months (95%CI: 27–54) and 34 months (95%CI: 25–41) for SR and TACE respectively (p = 0.30); while it was 25 months (95%CI: 9–43) for SR and 30 months (95%CI: 21–50) for TACE within B3-4 group. The median DFS was significantly different in the B1 subgroup: 13 months (95%CI: 8–17) for SR and 11 months (95%CI: 9–13) for TACE (p < 0.001). The B2 and B3-4 patients showed no significant difference between groups for DFS.

Secondary analysis: single HCC > 5 cm versus MC-IN tumors

To better understand the impact of surgery in the MC-Out cohort, we analysed as a secondary analysis, the OS by comparing MC-Out patients with MC-In patients in which surgery is the standard treatment. The baseline characteristics of the two cohorts are summarized in the Supplementary table 1. Seventy-four patients beyond MC had single nodule larger than 5 cm in size. Of them, 64 were treated by surgery (SR-SN). The median OS for SR-SN was 50.6 months (95%CI: 19.9–81.3), while it was 61.8 months (95% CI: 48.2–75.1) for the MC-IN group. At log-rank test, survival time was comparable between SR-SN and MC-IN group (p = 0.190). The median DFS was 18.8 months (95%CI: 10.3–27.3), 22.7 months (95%CI: 26.7–28.5) for SR-SN and MC-IN respectively. At comparison, SR-SN survival was similar to MC-IN (p = 0.954). Comparisons are depicted in Fig. 3.

Secondary analysis: multinodular HCC beyond MC versus MC-IN tumors

One-hundred and fifty patients had a multinodular disease beyond MC (>1 nodule >3 cm in size). Of them, 53 received surgery (SR-MN). They were first compared with the MC-IN

Table 2 Distribution comparison of the predictors between treatments in the original and weighted samples

Factors	Original sample N = 210[a]			Weighted sample N = 415.49		
	Surgery N = 113[a]	Tace N = 97[a]	Standardized difference	Surgery N = 210.62	Tace N = 204.87	Standardized difference
Cirrhosis, N (%)	83 (73.5)	89 (91.8)	−0.497	173.14 (82.2)	176.39 (86.1)	−0.107
MELD, mean (SD)	8.61 (2.98)	9.68 (2.48)	−0.390	9.08 (2.88)	9.10 (2.34)	−0.006
BILTOT, mean (SD)	0.96 (0.56)	1.22 (0.80)	−0.385	1.06 (0.60)	1.09 (0.71)	−0.057
INR, mean (SD)	1.14 (0.25)	1.18 (0.15)	−0.198	1.16 (0.23)	1.15 (0.14)	0.036
PLT, mean (SD)	171.81 (78.85)	110.52 (48.79)	0.935	143.66 (73.58)	136.10 (60.07)	0.112
N of Nodules >1, N (%)	51 (45.1)	86 (88.7)	−1.043	137.43 (65.25)	134.28 (65.54)	−0.006
Size >5 cm, N (%)	82 (72.6)	25 (25.8)	−1.059	106.27 (50.46)	100.73 (49.17)	0.026

MELD: model for end stage liver disease; BIL TOT: total bilirubin; INR: International Normalized Ratio; PLT: platelet count.
[a] The analysis includes only patients with complete data for all factors used in the model to estimate the inverse probability weights. This required the reduction of the sample size before matching.

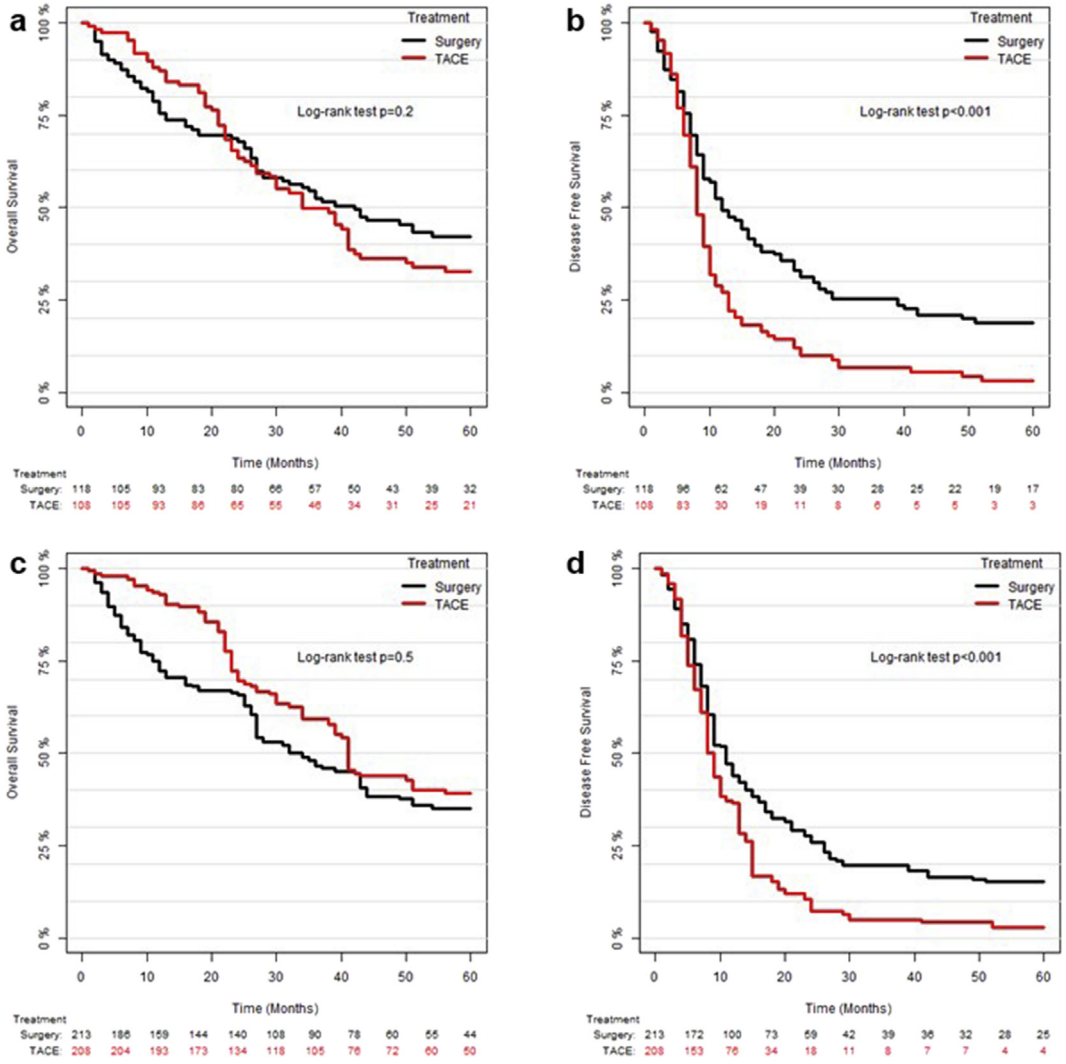

Figure 2 Survival curves before and after inverse probability weighting (IPW) for surgery and TACE. a) Overall survival before IPW; b) Disease-free survival before IPW c) OS after IPW d) DFS after IPW

group (n = 238). The median OS for SR-MN was 34.3 months (95%CI: 21.3–47.3) and 62 months (95%CI: 49.5–74.4) for MC-IN group. Significant differences were found between SR-MN versus MC-IN (p = 0.001). Surgically resected patients and beyond MC had a median DFS of 8.2 months (95%CI: 4.4-1.02), while MC-IN patients had a median DFS of 23.0 months (95%CI: 17.9–28.1; p < 0.001).

The SR-MN group was compared with the multinodular MC-IN proportion (MC-IN-MN; 2–3 nodules less than 3 cm without macrovascular invasion, n = 39). The median number of nodules for SR-MN was 2 (IQR: 2–3) with a median size of 5 cm (IQR: 4–6). The median number of nodules for MC-IN-MN was 2 (IQR: 2-2) with a median size of 2.2 cm (IQR 1.7–2.6). The median OS was 34.3 months (95%CI: 21.3–47.3) and 51.8 months (95%CI: 39.4–64.2) for SR-MN and MC-IN-MN respectively (p = 0.072). The median DFS was 8.2 months (95%CI: 4.4–12) for SR-MN and 18.2 months (95%CI: 3.3–33) for MC-IN-MN (p = 0.111). Comparison and overall results are depicted in Fig. 3.

Discussion

The present findings suggest that, in patients suffering from HCC beyond MC, the rate and median long-term survival were not significantly different between the two treatments, while a potentially more radical approach to the disease, such as the surgical resection, offers, in selected patients, some benefits in terms of disease-free survival when compared to chemoembolization.

Table 3 Uni- and multivariate Cox analyses on risk factors predicting overall survival (OS) and disease-free survival (DFS)

	Hazard ratios (95% CI)			
	OS, univariate	**OS, multivariate**	**DFS, univariate**	**DFS, multivariate**
TACE (vs. surgery)	1.218 (0.885; 1.676) p = 0.226	0.914 (0.615; 1.358) p = 0.657	1.791 (1.349; 2.378) p < 0.001	1.535 (1.085; 2.171) p = 0.015
Age (per year of increase)	1.008 (0.994; 1.023) p = 0.273	1.172 (1.000; 1.373) p = 0.050	0.991 (0.979; 1.004) p = 0.196	0.95 (0.83; 1.088) p = 0.458
Female (vs. male)	0.932 (0.627; 1.383) p = 0.725	0.884 (0.587; 1.331) p = 0.554	0.937 (0.67; 1.312) p = 0.706	0.96 (0.68; 1.353) p = 0.814
Cirrhosis (vs. no)	1.344 (0.875; 2.064) p = 0.177	1.073 (0.660; 1.743) p = 0.777	1.177 (0.82; 1.69) p = 0.377	0.784 (0.516; 1.189) p = 0.252
Child B (vs. A)	2.381 (1.595; 3.552) p < 0.001	2.219 (1.379; 3.572) p = 0.001	1.988 (1.394; 2.836) p < 0.001	1.822 (1.182; 2.808) p = 0.007
MELD (per unit)	1.077 (1.021; 1.136) p = 0.006	1.029 (0.962; 1.100) p = 0.407	1.062 (1.011; 1.114) p = 0.015	0.998 (0.937; 1.064) p = 0.962
HCV positive (vs. negative)	1.309 (0.953; 1.798) p = 0.096		1.26 (0.958; 1.657) p = 0.099	
HBV positive (vs. negative)	0.775 (0.514; 1.17) p = 0.225		0.845 (0.602; 1.188) p = 0.333	
Number of nodules > 1 (vs. ≤ 1)	1.356 (0.96; 1.917) 0.084	1.111 (0.650; 1.899) p = 0.700	1.789 (1.316; 2.432) p < 0.001	1.348 (0.855; 2.126) p = 0.199
Tumor size > 5 cm (vs. ≤ 5)	0.727 (0.526; 1.006) p = 0.055	0.797 (0.477; 1.330) p = 0.385	0.629 (0.472; 0.838) p = 0.002	0.893 (0.585; 1.365) p = 0.602

MELD: model for end stage liver disease; HCV: hepatitis C virus; HPB: hepatitis B virus.

As expected, the two treatment groups were not comparable for baseline characteristics. In particular, patients receiving TACE had more frequently a bilobar disease, a more deteriorated underling liver function, a multinodular presentation, and smaller nodules than patients who received surgery. In contrast, this latter group had a slide, but significant, poorer performance status and a higher rate of comorbid conditions. Overall, our treatment allocation mirrored other reports available in literature,[2,16] and it represents an evident confounder in the endeavour of comparing the two treatments for patients outside the Milan criteria. Noteworthy insights on the optimal approach for this cohort can be obtained only by randomized trials, but statistical methods such as the inverse probability weighting can lessen the role of confounders, as liver function and tumor burden, on long-term outcomes. After weighting, the two pseudo-populations were comparable for tumor burden and underlying liver function, but the estimates of overall survival rates did not substantially change. In addition, the results of the multivariate analysis confirmed that the type of treatment did not significantly affect the overall survival, but the underlying liver function, as measured by the Child classification, was an important variable in adjusting the risk of death. Therefore, we may speculate that, in most of the cases, the tumor was not the primary cause of decease even though we did not collect tumor-specific survival data. Age was another risk factors for overall survival. This result was somehow predictable since life expectancy decreases with age progression but it may also reflect the impact of age-related frailty in rescuing from long-term side effects of therapies,[17] since age-related decline in physiological reserve and functional capacity are inevitable and may negatively affect cure tolerance. The present results endorse previous memorandum on the key role of ageing in determining treatment long-term outcomes[18,19] and should motivate a more careful selection among elderly subjects.

Despite our findings are consistent with some reports,[20] and dissimilar from others in terms of overall survival,[21–25] the appraisal is challenging to achieve because of the different classification systems used. All the above publications used the BCLC staging system and compared the effect of treatments in patients classified as intermediate stage B which suffers high heterogeneity.[10] Moreover, high inconsistency in outcome measures[22,26–28] were reported when different managements were compared for the cure of single large tumors (>5 cm) or multinodular disease at presentation. Single large tumors are considered a borderline population, with conflicting results on prognosis. This observation induced to classify these patients as at intermediate stage,[29] or at least between early and intermediate stage (AB stage).[30] The present results fully support the latest EASL recommendation[7] stating that surgical resection is an appropriate option for single HCC of any size when hepatic function is preserved, and sufficient remnant liver volume is guaranteed postoperatively. In fact, we observed a similar median survival when the SR group with a single large tumor beyond the Milan criteria was compared to the early-stage (MC-IN) group.

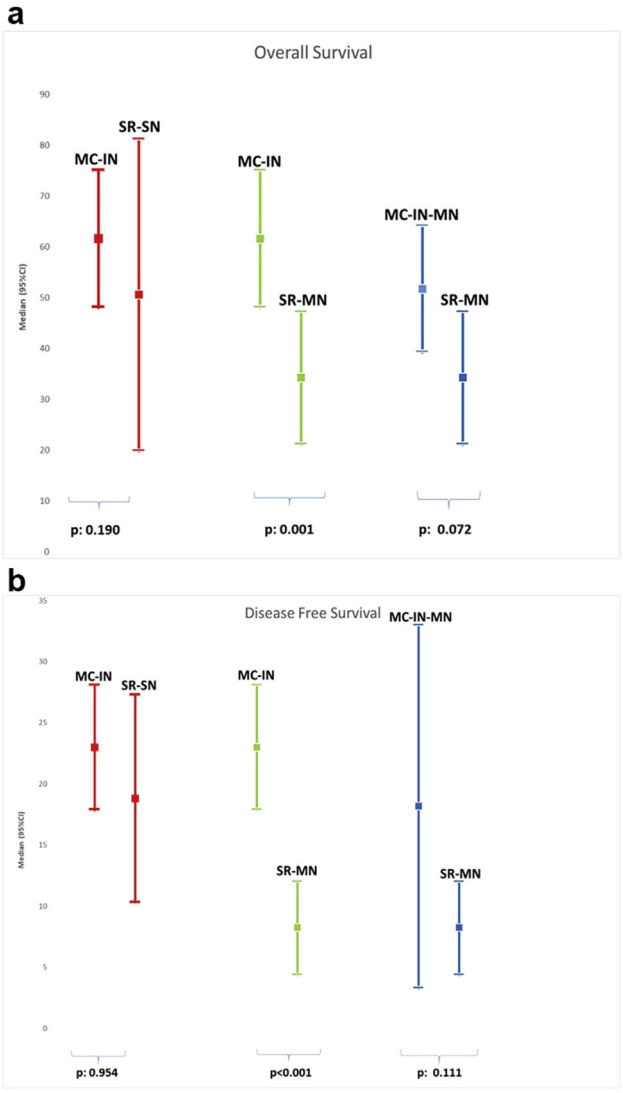

Figure 3 Box plots for a) median overall survival (OS) and b) disease-free survival (DFS) in different subgroups according to the number of nodules, size and staging, estimated by pairwise log-rank test. MC-IN: Milan Criteria-IN group; SR-SN: milan OUT surgical resection - single nodule; SR-MN: milan Out surgical resection - multinodular; MC-IN-MN: Milan Criteria-IN - Multinodular tumors

However, it should be acknowledged that the criteria for a safe operation (technical or general health conditions) may suffer from subjective judgement and local policies.

By analysing the population with multinodular disease, we observed that the surgical group, as compared with the early stage group (MC-IN), had a significant poorer median survival.. However, when surgical patients beyond MC and having a multinodular disease were compared with patients having a multinodular disease and classified as Milan-IN (2–3 nodules less than 3 cm in size), there was no significant difference in terms of survival. A RCT[4] investigating the role of surgery versus chemo-embolization, showed a survival benefit in patients who underwent surgery for multinodular HCC beyond MC. Our results are in line, since surgery allows a comparable survival rate with a reference population (MC-In patients) in which surgery is the standard of care. According to these results, surgery might be reconsidered as a treatment option also in multinodular disease and in cases of MC-out patients. However, these results should be considered with caution for the risk of type II error due to the small number of patients.

To better stratify diseases with a more aggressive behaviour from the ones with favourable prognosis, Bolondi *et al.*[10] proposed a sub-classification that takes into account the tumor burden along with the underlying liver function. By adopting this proposal, we found that surgery provided significant benefits in terms of disease-free survival over TACE but only for patients belonging to the B1 cluster. This subgroup comprises patients with high tumor burden but adequate liver function. The advantage of surgical resection may be due to the risk reduction of local and metastatic relapse but also to the lower risk of de-novo recurrence associated to advance liver dysfunction.

In general, the present findings suggest that surgical resection offered a significant benefit on disease-free survival. This gain was seen in both the observed and weighted cohorts. When an operation is feasible, liver resection appeared to guarantee a better disease control than TACE, even in neoplasm with more aggressive features than the tumors which are considered in an early stage presentation. In fact, our results showed that chemoembolization doubles the risk of tumor relapse than surgery, probably because of the reduced ability to control the intra-hepatic disease extension and so increasing the risk of metastatic recurrence.

Disease-free survival is an end-point for an increasing number of clinical studies of cancer treatment and can be considered both as a surrogate end-point and as an end-point in itself.[31] The period of time without relapse directly translates into more time that patients experience without cancer and cancer directed therapy. If a treatment would not make a person live any longer overall, but would make him live longer without documented disease, this may be an intrinsic added value of a treatment on quality of life. Moreover, disease relapse carries considerable health-associated and social costs for the time and medical resources required for recurrence treatments.

Several study limitations should be acknowledged. The major drawback is treatment allocation. This was an inevitable consequence of a subjective clinical decision based on tumor and patient characteristics. To limit this intrinsic bias of all retrospective studies, we adopted the inverse probability weighting approach to minimize the potential selection bias in analysing the association between treatments and outcomes. Second, excessive subgrouping could have produced type-II errors generating unreliable results. Third, tumor-specific survival could have been a more appropriate outcome measure to weigh the effect of treatments.

Conclusion

Surgical resection allowed a better control than TACE in patient bearing HCC beyond MC without macrovascular invasion. This translated into a significant benefit in terms of disease-free survival but not overall survival. Underlying liver function seemed the more important risk factor affecting overall survival. When the aggressiveness of an HCC is evaluated according to the tumor burden (as per the MC criteria), surgical resection is an appropriate option for single HCC of any size when hepatic function is preserved, while additional research is needed to address the best therapeutic options for multinodular disease exceeding the MC criteria, because the actual guideline indication may not be the best option for these patients.

Financial disclosure

No financial interest was disclosed by the authors.

Conflicts of interest

None declared.

References

1. Mazzaferro V, Regalia E, Doci R, Andreola S, Pulvirenti A, Bozzetti F *et al*. (1996) Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med* 334:693–699.
2. Hsu C-Y, Hsia C-Y, Huang Y-H, Su C-W, Lin H-C, Pai J-T *et al*. (2012) Comparison of surgical resection and transarterial chemoembolization for hepatocellular carcinoma beyond the Milan criteria: a propensity score analysis. *Ann Surg Oncol* 19:842–849.
3. Yang J, Yang J-Y, Yan L-N, Wen T-F, Li B, Wang W-T. (2018) Hepatic resection vs. transarterial chemoembolization for hepatocellular carcinoma beyond the Milan criteria with portal hypertension. *Dig Liver Dis* 50:713–719.
4. Yin L, Li H, Li A-J, Lau WY, Pan Z-Y, Lai ECH *et al*. (2014) Partial hepatectomy vs. transcatheter arterial chemoembolization for resectable multiple hepatocellular carcinoma beyond Milan Criteria: a RCT. *J Hepatol* 61:82–88.
5. Clavien P-A, Lesurtel M, Bossuyt PMM, Gores GJ, Langer B, Perrier A, OLT for HCC Consensus Group. (2012) Recommendations for liver transplantation for hepatocellular carcinoma: an international consensus conference report. *Lancet Oncol* 13:e11–e22.
6. Mazzaferro V, Bhoori S, Sposito C, Bongini M, Langer M, Miceli R *et al*. (2011) Milan criteria in liver transplantation for hepatocellular carcinoma: an evidence-based analysis of 15 years of experience. *Liver Transplant* 17(Suppl. 2):S44–S57.
7. Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul J-L *et al*. (2018) EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 69:182–236.
8. Torzilli G, Belghiti J, Kokudo N, Takayama T, Capussotti L, Nuzzo G *et al*. (2013) A snapshot of the effective indications and results of surgery for hepatocellular carcinoma in tertiary referral centers: is it adherent to the EASL/AASLD recommendations?: an observational study of the HCC East-West study group. *Ann Surg* 257:929–937.
9. Liang L, Xing H, Zhang H, Zhong J, Li C, Lau WY *et al*. (2018) Surgical resection versus transarterial chemoembolization for BCLC intermediate stage hepatocellular carcinoma: a systematic review and meta-analysis. *HPB* 20:110–119.
10. Bolondi L, Burroughs A, Dufour J-F, Galle PR, Mazzaferro V, Piscaglia F *et al*., for the STROBE Initiative. (2012) Heterogeneity of patients with intermediate (BCLC B) Hepatocellular Carcinoma: proposal for a subclassification to facilitate treatment decisions. *Semin Liver Dis* 32: 348–359.
11. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. (2007) The strengthening the reporting of observational studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 4: e296.
12. Famularo S, Di Sandro S, Giani A, Lauterio A, Sandini M, De Carlis R *et al*. (2018) Long-term oncologic results of anatomic vs. parenchyma-sparing resection for hepatocellular carcinoma. A propensity score-matching analysis. *Eur J Surg Oncol* 44:1580–1587.
13. Famularo S, Giani A, Di Sandro S, Sandini M, Giacomoni A, Pinotti E *et al*. (2018) Does the Pringle maneuver affect survival and recurrence following surgical resection for hepatocellular carcinoma? A western series of 441 patients. *J Surg Oncol* 117:198–206.
14. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET *et al*. (1982) Toxicity and response criteria of the eastern cooperative Oncology group. *Am J Clin Oncol* 5:649–656.
15. Famularo S, Di Sandro S, Giani A, Lauterio A, Sandini M, De Carlis R *et al*. (2018) Recurrence patterns after anatomic or parenchyma-sparing liver resection for hepatocarcinoma in a western population of cirrhotic patients. *Ann Surg Oncol* 25:3974–3981.
16. Cucchetti A, Mazzaferro V, Pinna AD, Sposito C, Golfieri R, Serra C *et al*. (2017) Average treatment effect of hepatic resection versus locoregional therapies for hepatocellular carcinoma. *Br J Surg* 104:1704–1712.
17. Sandini M, Pinotti E, Persico I, Picone D, Bellelli G, Gianotti L. (2017) Systematic review and meta-analysis of frailty as a predictor of morbidity and mortality after major abdominal surgery. *BJS Open* 1: 128–137.
18. Famularo S, Di Sandro S, Giani A, Angrisani M, Lauterio A, Romano F *et al*. (2019) The impact of age and aging on hepatocarcinoma surgery: short- and long-term outcomes in a multicenter propensity-matched cohort. *Liver Int*. https://doi.org/10.1111/liv.14075.
19. Kaibori M, Yoshii K, Yokota I, Hasegawa K, Nagashima F, Kubo S *et al*. (2019) Impact of advanced age on survival in patients undergoing resection of hepatocellular carcinoma: report of a Japanese nationwide survey. *Ann Surg* 269:692–699.
20. Tian X, Dai Y, Wang D-Q, Zhang L, Sui C-G, Meng F-D *et al*. (2015) Transarterial chemoembolization versus hepatic resection in hepatocellular carcinoma treatment: a meta-analysis. *Drug Des Dev Ther* 9: 4431–4440.
21. Pecorelli A, Lenzi B, Gramenzi A, Garuti F, Farinati F, Giannini EG *et al*. (2017) Curative therapies are superior to standard of care (transarterial chemoembolization) for intermediate stage hepatocellular carcinoma. *Liver Int* 37:423–433.
22. Zhong J-H, Wu F-X, Li H. (2014) Hepatic resection associated with good survival for selected patients with multinodular hepatocellular carcinoma. *Tumor Biol* 35:8355–8358.
23. Qi X, Wang D, Su C, Li H, Guo X. (2015) Hepatic resection versus transarterial chemoembolization for the initial treatment of hepatocellular carcinoma: a systematic review and meta-analysis. *Oncotarget* 6.
24. Oliveri RS, Wetterslev J, Gluud C. (2011) Transarterial (chemo)embolisation for unresectable hepatocellular carcinoma. *Cochrane Database*

Syst Rev, CD004787. https://doi.org/10.1002/14651858.CD004787.pub2.

25. Liu W, Zhou J-G, Sun Y, Zhang L, Xing B-C. (2015) Hepatic resection improved the long-term survival of patients with BCLC stage B hepatocellular carcinoma in asia: a systematic review and meta-analysis. *J Gastrointest Surg* 19:1271–1280.

26. Yang X-D, Pan L-H, Wang L, Ke Y, Cao J, Yang C *et al*. (2015) Systematic review of single large and/or multinodular hepatocellular carcinoma: surgical resection improves survival. *Asian Pac J Cancer Prev* 16:5541–5547.

27. Deng L, Yang C, Li L-Q, Zhong J-H. (2015) Hepatic resection improves long-term survival of patients with large and/or multinodular hepatocellular carcinoma. *J Gastrointest Surg* 19:2288–2289.

28. Stevens CL, Awad A, Abbas SM, Watters DAK. (2017) Systematic review and meta-analysis of hepatic resection versus transarterial chemoembolization for solitary large hepatocellular carcinoma. *HPB* 19:653–658.

29. Yau T, Tang VYF, Yao T-J, Fan S-T, Lo C-M, Poon RTP. (2014) Development of Hong Kong liver cancer staging system with treatment stratification for patients with hepatocellular carcinoma. *Gastroenterology* 146:1691–1700.e3.

30. Vitale A, Burra P, Frigo AC, Trevisani F, Farinati F, Spolverato G *et al*. (2015) Survival benefit of liver resection for patients with hepatocellular carcinoma across different Barcelona Clinic Liver Cancer stages: a multicentre study. *J Hepatol* 62:617–624.

31. Robinson AG, Booth CM, Eisenhauer EA. (2014) Disease-free survival as an end-point in the treatment of solid tumours – perspectives from clinical trials and clinical practice. *Eur J Cancer* 50:2298–2302.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.hpb.2019.12.011.