STA490: Statistical Consulting

Prof. Dr. Leonhard Held

EBPI, Department of Biostatistics

Benefit of prolonged and earlier treatment of Hepatocellular carcinoma patients with TACE

Analysis for Dr. Joachim Mertens and Anouk Zgraggen, Department of Gastroenterology and Hepatology of the University Hospital Zurich

Supervision by Monika Hebeisen

Joël Meili (*joel.meili@uzh.ch*)

Version of January 12, 2021

1 Abstract

This study aims to investigate if there is a benefit of prolonged and earlier treatment of hepatocellular carcinoma (HCC) patients with TACE. Patients were seen regularly for at least five years. The methods used were Kaplan-Meier plots, competing risk analysis and Cox-regression. The analysis is subject to the assumption that the patients in BCLC stages 0-A and C are not representative as patients are probably more/less ill than the average patient in said stage. The analysis showed that there is a tendency for improved overall survival for patients in BCLC stages B and C. Contrarily, TACE did not improve the overall survival for patients of the earlier stages 0 and A. The comparison of overall and progression-free survival curves between the different BCLC stages showed that there is no evidence for a difference between patients in stages 0-A and B. Competing risk analysis showed that patients regardless of their staging die quickly from hepatic failure compared to other causes of death. Confounders such as sex, age and underlying conditions did not influence the effect of the BCLC stages on the overall survival. Based on the results we would recommend only patients of BCLC stages B and C to receive earlier and prolonged treatment with TACE.

2 Introduction

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer. It makes up to a third of all cancer deaths in the western world. There are various factors that lead to liver cirrhosis which can trigger the development of HCC such as hepatitis B and C, chronic alcohol consumption and age. The staging of the cancer is important as it indicates the recommended first-line treatment. Currently there are no curative treatments available for intermediate to terminal stages. HCC as many other forms of cancer is oligosymptomatic, especially in early stages, which makes early diagnosis a challenging task. Consequentially, advanced stages account for up to 70% of the newly diagnosed cases of HCC.

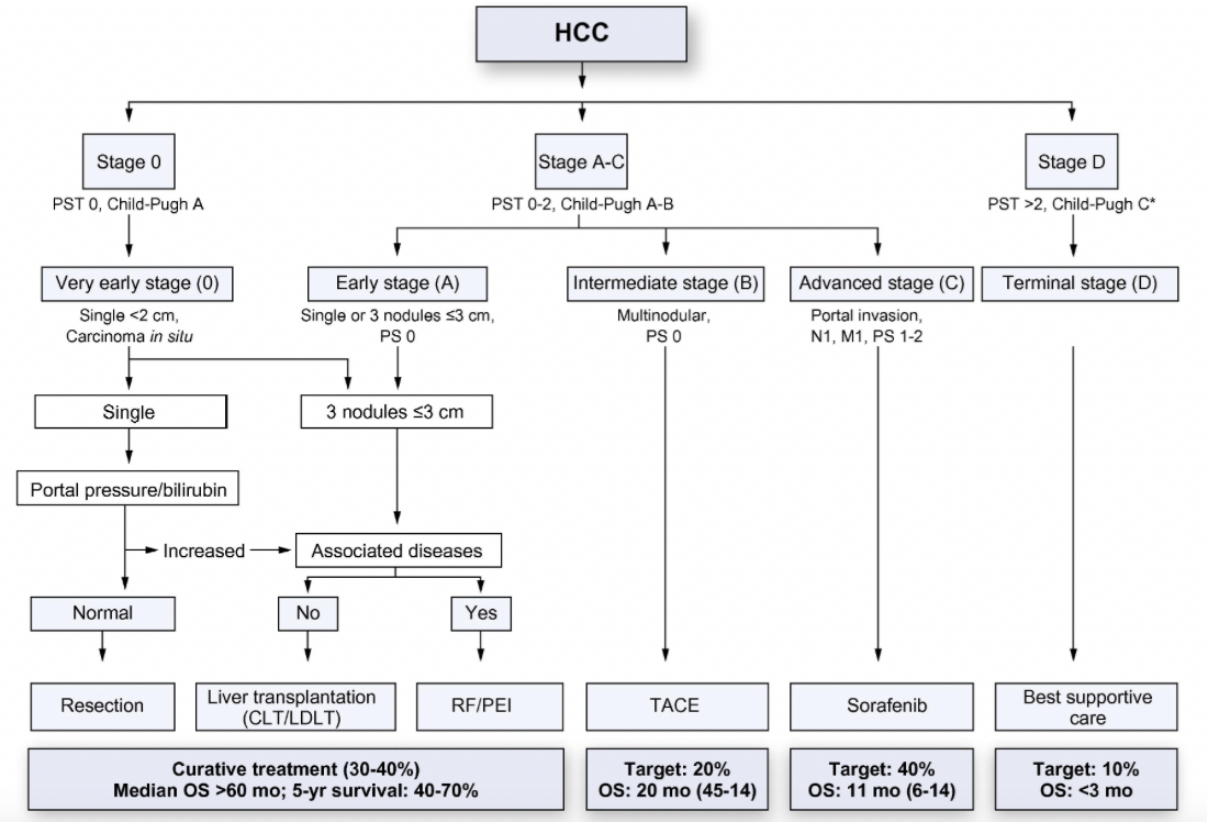

Figure 1: BCLC staging decision flowchart Llovet et al. (2012)

Figure 1 illustrates how HCC cases are assessed by the BCLC (Barcelona Clinic Liver Cancer) staging system and the recommended first-line treatment for each stage. From the figure it can be seen that only for patients of the earliest stage 0 it is recommended to do resection of the tumour. The BCLC guidelines suggest that only patients with BCLC stage B should be treated with TACE. TACE (transarterial chemoembolization) is a localized cancer therapy where the chemotherapeutic agent is delivered directly into the vessel which supplies the tumour. Usually, TACE is very well tolerated by the patients, however guidelines suggest not more than two TACE treatments per patient. Therefore, this study aims to investigate whether additional TACE treatments given to patients regardless of their initial staging improve survival compared to guidelines.

3 Research Questions

1. Does prolonged and earlier treatment of patients with HCC with TACE lead to improved overall survival compared to guidelines?
2. Is there a difference in progression-free survival between patients in different BCLC stages?
3. Do causes of death vary between the patients in different BCLC stages?
4. Is there an influence of patient characteristics on the overall survival of HCC patients?

4 Methods

Study Design

Type of study

This is an observational study with retrospectively collected data. Patients were followed-up every three months after TACE for a period of at least five years. If patients were to be tumour-free after two years they were seen every six months.

Study population

In this study the population consisted of patients who were diagnosed with HCC irrespective of cause and treated with TACE. All patients had access to standard care and the diagnosis was made not more than ten years ago. The HCC diagnosis was made according to the EASL-EORTC criteria Llovet et al. (2012) and was supported by either sonography, CT, MRI or biopsy. All patients were at least 18 years or older. The study includes all patients first treated with TACE at USZ between January 2010 and December 2019.

Data collection

The data contains an initial assessment of the patient's baseline values before the first TACE treatment was to be administered. This included the staging of the BCLC, counting the number of lesions present in the liver, measuring the lesion with the largest diameter and whether the HCC had already spread or invaded the neighboring blood vessels. Further it was assessed whether the patient had any underlying conditions such as hepatitis B or C and if the patient was treated for those prior to the study. During the regular follow-ups information was updated on whether the patient had any therapy since the last follow-up e.g. resection of the liver, transplantation or TACE. Further, it was assessed whether the disease had progressed since the last follow-up. Progression was defined as whether the tumour had grown locally, had spread or had invaded neighbouring blood vessels. Assessing whether progression had occurred was important as it influenced the following treatment plan for that patient as TACE was usually not given to a patient after progression. However, there was no clear protocol and the decision to go for another round of TACE was made individually for each patient.

Variables

Table 1: Variables and their definition

Name	Definition
Sex	Sex of the patient
Age	Age of patient at 2019-31-12
DoBIRTH	Date of birth
DoHCC	Date of HCC diagnosis
DLS	Date when patient was last seen
TACE1	Date of TACE1
TACE2	Date of TACE2
TACE3	Date of TACE3
TACE4	Date of TACE4
TACE5	Date of TACE5
TACE6	Date of TACE6
TACE7	Date of TACE7
DISEASE	Underlying disease
Dead	Patient is dead
DoDEATH	Date of death
CAUSE	Cause of death
PROGRESS	Tumour has progressed after TACE
DoP	Date of progression
BCLC	BCLC staging of tumour before TACE
Resection	Tumour was surgically removed before first TACE
Transplantation	Liver was transplanted after first TACE
RFA	Patient received RFA treatment before first TACE
MWA	Patient received MWA treatment before first TACE
METASTASIS	Had the tumour spread before first TACE
ANGIONINVASION	Had the tumour invaded neighbouring blood vessels before first TACE
NODULES #	Number of lesions before first TACE
NODULEmax	Diameter of largest lesion
Biopsy	Patient had biopsy before first TACE

Primary and secondary outcomes

Primary outcome:

1. Median overall survival (OS) compared to guidelines for all stages and 5-year OS for early stages (0-A)

Secondary outcomes:

2. Progression-free survival (PFS) for all stages
3. Competing risk plot for death due to different causes
4. Overall survival of different stages after adjustment for baseline characteristics and transplantation as a time-varying covariate

Statistical Analysis

Data Preparation

The data set was already quite clean so there was not much data cleaning left to do. First, I went through the initial list of variables and selected those relevant to the analysis. Some variables could not be used in the analysis as they were maximum or minimum values that were observed in the whole follow-up period and did not have a time stamp. Next, I created an arbitrary ID variable to be able to identify the patients later on. As there were many underlying diseases that caused HCC in the patients I looked at those that were most common and coded the rest as other. I used `grepl` to extract the date of transplantation out of the original transplantation variable. Further, I calculated the difference between dates such as difference between date of death and date of first TACE to prepare the necessary variables for the survival analysis. I chose to divide the difference in days by 30.4 to get the difference in months between two events as the available data from the BCLC guidelines paper Llovet et al. (2012) was in months as well.

Descriptive Statistics and Simple Methods

Patient characteristics were displayed by BCLC stage. Numerical data was shown with mean and standard deviation, categorical with numbers and percentages. Further, I chose standard bar charts to visualize proportions e.g. sex of the patient or the BCLC stage.

Visualization Methods

For the visualization of the overall and progression-free survival I used Kaplan-Meier plots Kaplan and Meier (1958). Kaplan-Meier can be used to show the estimated survivor function for our data. In survival analysis the survivor function gives the probability of not encountering the event e.g. death just before some time t.

I visualized the results of the adjusted Cox-regression with a Forest plot with 95%-confidence intervals for the estimated effects. This version of the Forest plot also includes the number of observations within a given subset. Further, I used the scaled Schoenfeld residuals to check whether there was evidence for a violation of the proportional hazards assumption. If there was evidence against the proportional hazards assumption the plot would show a smoothed line that is changing over time - not horizontal.

Description of advanced statistical methods

I used the square-and-add method to calculate the difference in median OS between the estimated results and the BCLC guidelines paper Llovet et al. (2012).

To test whether the OS and PFS curves differ between different BCLC stages I used pairwise log-rank tests. A log-rank test Peto and Peto (1972) is a hypothesis test which compares two survival curves where the Null hypothesis states that there is no difference between the two curves. In case of multiple comparisons one has to consider that the probability for at least one significant result is increased the more tests are being made. Hence, I chose the Benjamini-Hochberg correction Benjamini and Hochberg (1995) to control for the family-wise error rate.

Competing risk analysis is used when patients can experience different type of events e.g. death from cardiovascular causes or death from HCC. I used competing risk analysis to calculate the cumulative incidence rate of causes of death for each BCLC stage to investigate whether there was a difference in causes of death in the different BCLC stages.

Cox-regression Cox (1972) is used to estimate the effects of one or more covariates on the time a specified event takes to happen. In this study I used Cox-regression to investigate the differences between the BCLC stages and possible confounders such as age on the OS. The estimated effects are hazard ratios. A hazard ratio is the ratio between hazard rates e.g. the ratio between the hazard rate of being male and the hazard rate of being female. Hazard is to be interpreted as an instantaneous risk of death. Hazard ratios smaller than one indicate lower hazard rates and therefore higher survival rates whereas hazard ratios larger than one indicate higher hazard rates and therefore lower survival rates. Further, I used 95%-confidence intervals to show whether there was evidence for an observed hazard ratio to be different from one. The estimates of the Cox-regression are calculated by maximizing the partial likelihood.

To investigate whether patient characteristics influence the OS time I used a proportional hazards model (Cox-regression). First, I identified a set of variables that could be possible confounders in this analysis. These are: sex, initial BCLC staging, underlying disease, age at first TACE treatment, time from diagnosis to first TACE treatment, whether the patient received RFA and/or MWA treatment and whether the patient had a liver transplantation. Second, I had to determine what time stamp should be used as t=0. In the previous analyses I used time from diagnosis to be able to compare the results to the BCLC guidelines paper Llovet et al. (2012). However, analysis with time from diagnosis has some drawbacks as immortal time bias Lévesque et al. (2010) is induced. Immortal time bias is caused in this case because in the period from diagnosis to the first TACE all patients are immortal. Patients who die between diagnosis and their (potential) first TACE treatment are not considered in the analysis. So the survival curves of time from diagnosis to death are biased to higher values because some patients who die early are not considered. I chose to settle on time from first TACE to avoid the effect of the immortal time bias. Third, I had to manage the issue of transplantation as the status of the variable changes over time. Hence, I chose to model transplantation as a time-varying variable where patients with an eventual transplantation have two observational periods: not transplanted, time to transplantation, no event / transplanted, time to event or censoring, censoring or event.

Implementation

All analyses were performed in the R programming language (R Core Team, 2018) using base packages and the following analysis-specific packages: `biostatUZH`, `tableone`, `survival` and `cmprsk`. Data analysis was done according to STROBE reporting guidelines Von Elm et al. (2007).

5 Results

Descriptive Statistics of Patient Baseline Characteristics

The data set consists of 267 patients and their patient characteristics. Two patients were removed from the analysis as one had a transplantation before being diagnosed with HCC and the other was removed because the date last seen was equal the date of first treatment with TACE. Hence, for this patient time to censoring was zero. Figure 2 shows that the sex of the study population is unbalanced as 81.9% of the patients are male. Additionally, 57.7% of the patients presented at least BCLC stage B at diagnosis, which means that over half of the patients were deemed incurable of the disease.

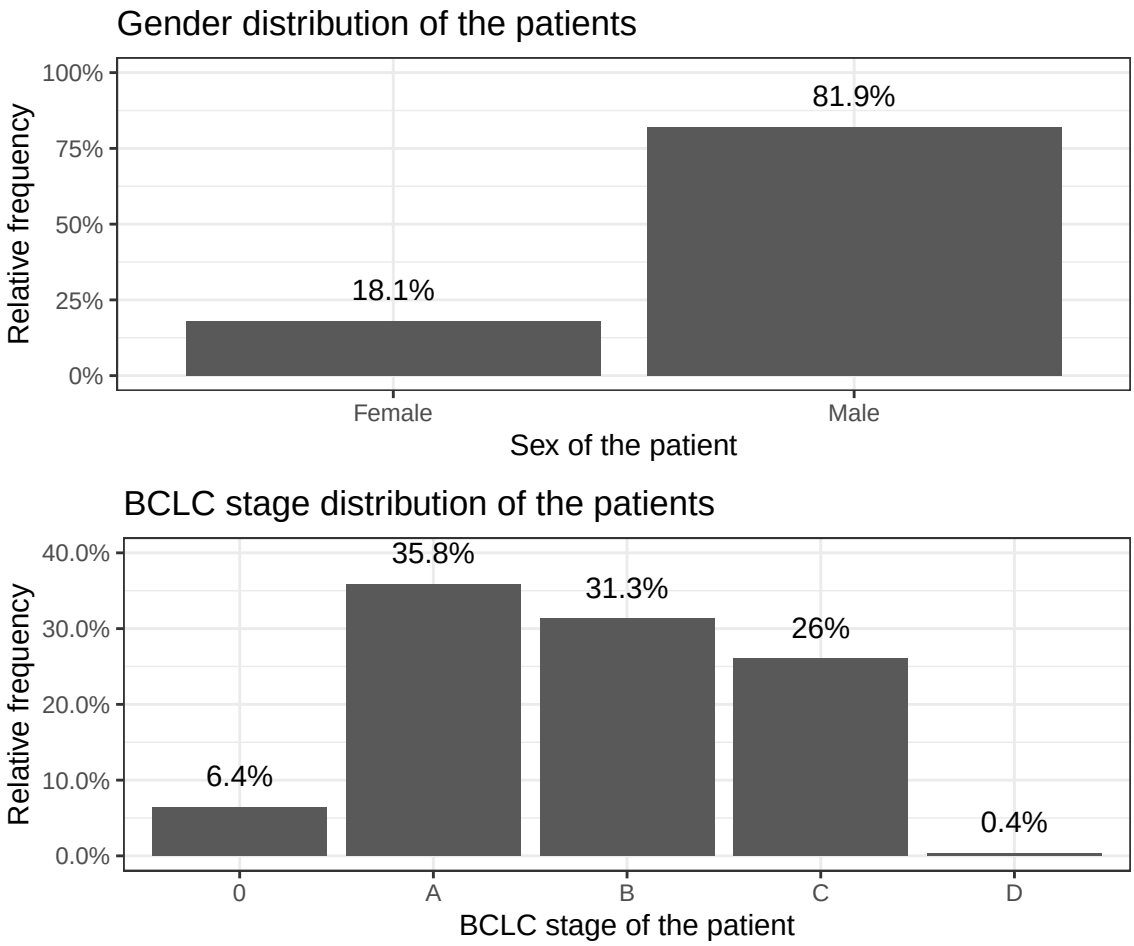

Figure 2: Observed distribution of gender and BCLC stage of the patients

Table 2 summarizes some of the patient characteristics at diagnosis. There were 265 patients in total: 17 with stage 0, 95 with stage A, 83 with stage B, 69 with stage C and one patient with stage D. For the following analyses the patients of groups 0-A were grouped together as one group (early stages) to ensure comparability to the BCLC guideline papers Llovet et al. (2012). Additionally, the sole patient with stage D was ommited from further analyses due to the lack of sample size available in that stage. The proportion of males in each stage is larger as compared to females, however is lowest in stage 0. The mean age at diagnosis is again lowest in stage 0 as compared to the other stages which could be due to fewer observations made. Otherwise the mean age at diagnosis is approximately the

same in all stages. As previously mentioned, there are various factors that can trigger the development of HCC. Alcohol however is the most common underlying cause here in all BCLC stages.

Table 2: Patient characteristics at diagnosis by BCLC stages (SD: standard deviation; HCC: hepatocellular carcinoma; HBV: hepatitis B virus; HCV: hepatitis C virus; NASH: non-alcoholic steatohepatitis)

	0	**A**	**B**	**C**	**D**
n	17	95	83	69	1
Sex = m (%)	11 (64.7)	78 (82.1)	68 (81.9)	59 (85.5)	1 (100.0)
Age at diagnosis (mean (SD))	61.94 (9.83)	64.61 (9.35)	65.67 (9.60)	64.74 (10.79)	65.00 (NA)
Underlying cause for HCC (%)					
Alcohol	8 (47.1)	30 (31.6)	27 (32.5)	21 (30.4)	0 (0.0)
HBV	0 (0.0)	10 (10.5)	5 (6.0)	6 (8.7)	0 (0.0)
HCV	4 (23.5)	22 (23.2)	24 (28.9)	11 (15.9)	0 (0.0)
kryptogen	0 (0.0)	9 (9.5)	1 (1.2)	11 (15.9)	0 (0.0)
NASH	0 (0.0)	7 (7.4)	9 (10.8)	4 (5.8)	0 (0.0)
Other	5 (29.4)	17 (17.9)	17 (20.5)	16 (23.2)	1 (100.0)

Table 3 shows additional patient characteristics that were observed during the follow-up period. First, the number of TACE treatments given to patients in BCLC stage B is higher than compared to the other stages. Adherence to guidelines indicates that a patient received two or less TACE treatments during follow-up. Our data shows that 92.0% of patients in stages 0-A did not receive more than two TACE treatments. Further, time from diagnosis to first TACE was the highest (6.66 months) for patients in BCLC stage B. Contarily, patients in stage B had the lowest time (4.51 months) from first TACE to second TACE. Our data indicates that a third of the patients in stages 0-A received a liver transplantation, compared to only a fifth in stage B and a tenth in stage C. The percentages of stage C patients presenting with metastases and/or angioinvasion was quite balanced. Further, the number of lesions was largest for patients in stage B. However, the size of the largest lesion was largest for patients in stage C.

Table 3: Additional patient characteristics by BCLC stages after omitting the sole patient with BCLC stage D and grouping patients of stages 0 and A as one group (SD: standard deviation; HCC: hepatocellular carcinoma; HBV: hepatitis B virus; HCV: hepatitis C virus; NASH: non-alcoholic steatohepatitis; Patient received other treatment additional to TACE: patient received RFA and/or MWA)

	0-A	**B**	**C**
n	112	83	69
Number of TACE treatments received (mean (SD))	1.77 (1.06)	2.20 (1.32)	1.99 (0.99)
Patient has received TACE treatment according to guidelines (2 or less) = Guideline (%)	92 (82.1)	56 (67.5)	50 (72.5)
Age at first TACE (mean (SD))	64.51 (9.46)	66.25 (9.40)	65.07 (10.77)
Time from diagnosis until first TACE (in months) (mean (SD))	4.71 (6.77)	6.66 (11.41)	5.14 (11.86)
Time from first TACE to second TACE (in months) (mean (SD))	8.16 (8.18)	4.51 (3.51)	4.60 (3.97)
Liver transplantation after TACE = yes (%)	38 (33.9)	18 (21.7)	9 (13.2)
Liver resection before TACE = yes (%)	9 (8.0)	4 (4.9)	6 (8.8)
Metastases before TACE = yes (%)	0 (0.0)	0 (0.0)	44 (63.8)
Angioinvasion before TACE = yes (%)	0 (0.0)	0 (0.0)	37 (53.6)
Biopsy before TACE = yes (%)	48 (42.9)	38 (45.8)	26 (37.7)
Competing risk event (%)			
No event	60 (53.6)	38 (45.8)	13 (18.8)
cardiovascular	1 (0.9)	2 (2.4)	0 (0.0)
HCC	4 (3.6)	11 (13.3)	23 (33.3)
hepatic failure	15 (13.4)	15 (18.1)	22 (31.9)
other	24 (21.4)	10 (12.0)	7 (10.1)
sepsis	8 (7.1)	7 (8.4)	4 (5.8)
Number of nodules/lesions before TACE (mean (SD))	1.63 (0.82)	3.23 (1.69)	2.99 (2.87)
Size of the largest nodule/lesion before TACE (mean (SD))	32.31 (20.72)	46.99 (24.41)	58.03 (35.88)
Patient has received other treatment additional to TACE = yes (%)	16 (14.3)	8 (9.6)	4 (5.9)

Median Overall Survival for Patients of All Stages

In this subsection I show the results of the survival analysis for the OS of the patients in all stages with the help of a Kaplan-Meier plot. Figure 3 shows the estimated survivor functions with 95%-confidence intervals for the OS in all BCLC stages. From the figure it can be seen that the OS is lowest for patients with BCLC stage C when compared to the other two stages. Further, survivor functions of stages 0-A and B cross after 50 months.

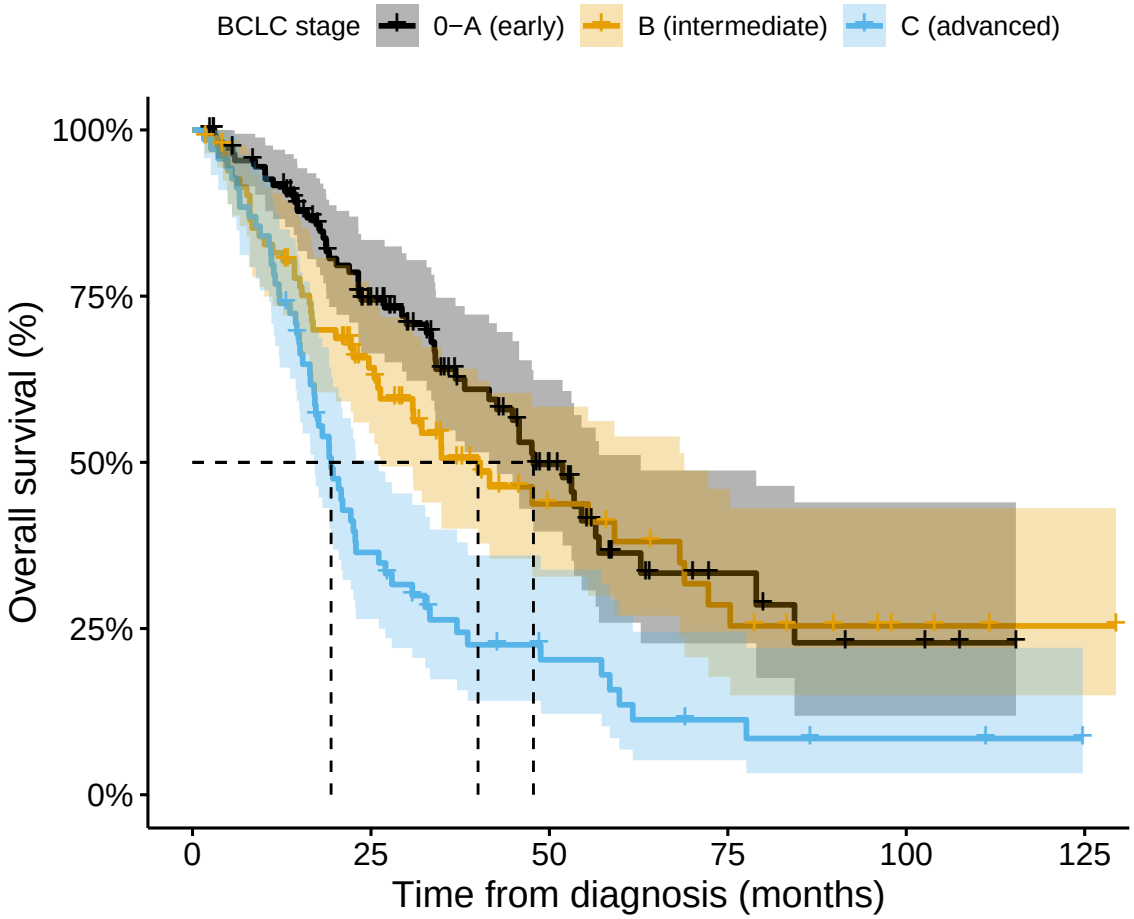

Figure 3: Kaplan-Meier curves for the overall survival of patients for all BCLC stages with 95%-confidence intervals visualized as coloured areas

Table 4 summarizes the estimated median OS times and their 95%-confidence intervals in comparison to available data from other studies such as the BCLC guidelines papers Llovet et al. (2012), Llovet et al. (2018) and Forner et al. (2018). The table shows that there is a tendency for a decreased OS for patients with stages 0-A when compared to the guidelines papers. Unfortunately, the mentioned studies did not provide confidence intervals. Therefore, the difference in median OS for stages 0-A cannot be calculated. Additionally, the 5-year OS is smaller as well when compared to Llovet et al. (2012). Based on our data patients in stage B have a tendency for an increased OS as the estimated difference in median OS is 20 months. However, the 95%-confidence interval for the difference is wide (from -8.5 to 49.5 months) and therefore we see no evidence for a difference in median OS. Contrarily, the difference in median OS for patients with stage C shows evidence for an increased OS (8.4 months, 95%-confidence interval from 4.6 to 16.8 months).

Table 4: Observed median overall survival time (in months) compared to available data from other studies for all BCLC stages and the observed 5-year overall survival compared to available data from other studies for BCLC stages 0-A

Study	**Median OS**	**95%-confidence interval**
Stages 0-A		
Observed data	47.8	from 41.5 to 62.8
Llovet et al. (2012)	≥60	-
Llovet et al. (2018)	≥60	-
Difference Observed and Llovet et al. (2012)	-	-
Stage B		
Observed data	40	from 26.3 to 68.9
Llovet et al. (2012)	20	from 14 to 45
Llovet et al. (2018)	≥30	-
Difference Observed and Llovet et al. (2012)	20	from -8.5 to 49.5
Stage C		
Observed data	19.4	from 17 to 26.1
Llovet et al. (2012)	11	from 6 to 14
Llovet et al. (2018)	≥10	-
Forner et al. (2018)	≥12	-
Difference Observed and Llovet et al. (2012)	8.4	from 4.6 to 16.8
Study	**5-year OS**	**95%-confidence interval**
Stages 0-A		
Observed data	36%	from 26% to 51%
Llovet et al. (2012)	-	from 40% to 70%

Further, I used pairwise log-rank tests with Benjamini-Hochberg correction for multiple testing to compare the OS between the different BCLC stages. Our data indicates that there is very strong evidence ($p < 0.0001$) for a difference between OS for patients in stages 0-A and C. Further, there is strong evidence for a difference in OS between stages B and C ($p = 0.002$). However, there seems to be no evidence for a difference in OS for patients in stages 0-A and B ($p = 0.27$).

Progression-free Survival for Patients of All Stages

PFS is defined as whether the disease had progressed or the patient had died. As previously mentioned, progression ocurred when the tumour had either grown locally, had spread or had invaded neighbouring blood vessels. Figure 4 shows the Kaplan-Meier plot for the PFS of patients for all BCLC stages. Similar to figure 3 patients in stage C had the worst outcome in PFS survival compared to the other two stages.

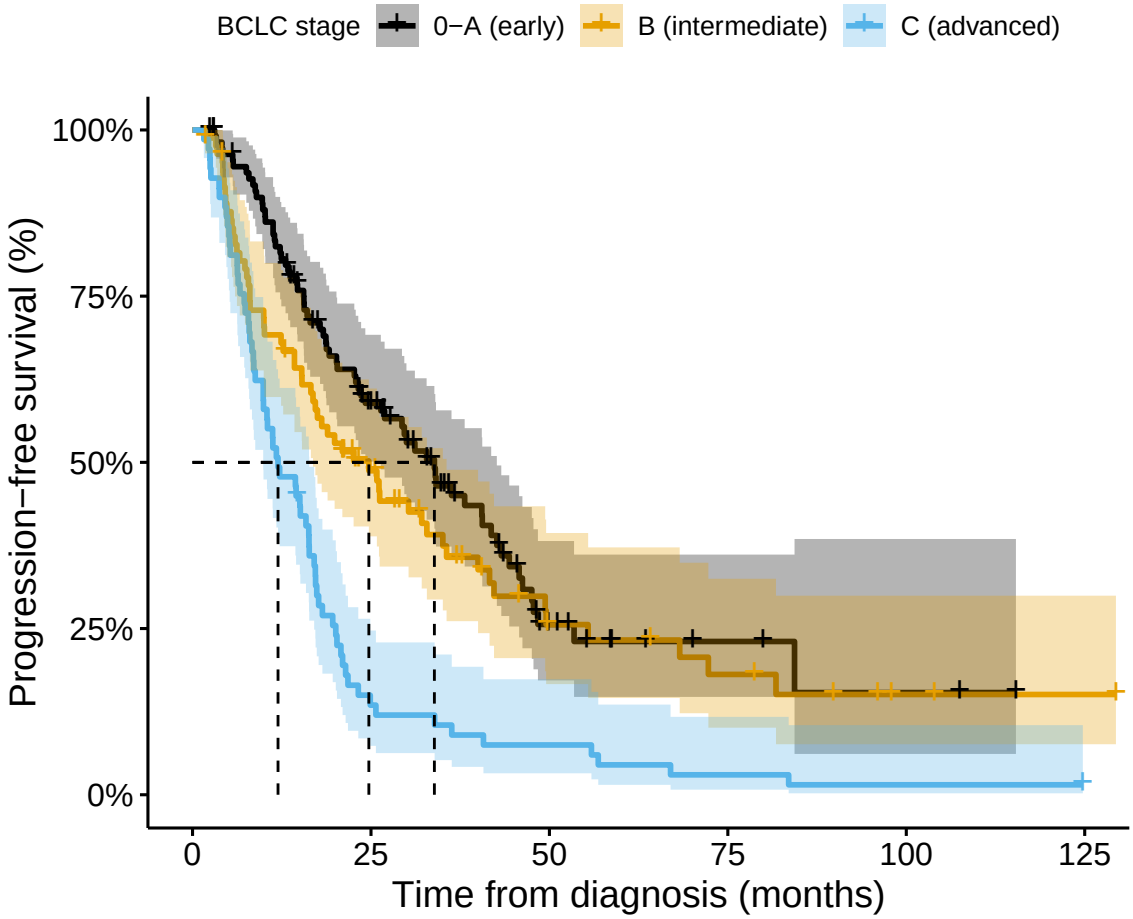

Figure 4: Kaplan-Meier curves for the progression-free survival of patients for all BCLC stages with 95%-confidence intervals visualized as coloured areas

First, patients in stages 0-A had an estimated median PFS of 33.9 months with a 95%-confidence interval from 26.4 to 42.6 months. Second, patients in stage B had an estimated median PFS of 24.7 months with a 95%-confidence interval from 16.9 to 35.6 months. Third, patients in stage C had an estimated median PFS of 12 months with a 95%-confidence interval from 9.9 to 16.3 months.

As there was no external data available from other studies I was not able to compare our estimated median PFS times to patients with other treatments. Thus, I used a pairwise log-rank test with Benjamini-Hochberg adjustment for multiple testing to compare the PFS between the BCLC stages. Based on the pairwise log-rank tests there is strong evidence for a difference in PFS for stage C and the other stages ($p < 0.0001$, $p < 0.0001$). However, there is no evidence for a difference in PFS between stages 0-A and B ($p = 0.19$). The results are similar to what was observed in the pairwise log-rank tests

for the OS, however the evidence for a difference between the patients in stages B and C is stronger for the PFS ($p < 0.0001$) than the OS ($p = 0.002$).

Competing Risk Plot for Death Due To Different Causes

Figure 5 shows the cumulative incidence curves of competing risks of death since the first TACE treatment. In the beginning patients with stages 0-A die mostly from hepatic failure. However, after some time most patients die from other causes unrelated to the liver. Additionally, when compared to the other stages only a small proportion of patients dies because of HCC. Patients in stage B mostly die from hepatic failure and HCC. Similarly to patients in stages 0-A patients in stage B die fairly quick from hepatic failure. Patients in stage C die almost exclusively from hepatic failure and HCC. Compared to the other stages patients in stage C not only die quickly due to hepatic failure but also due to HCC. Based on our data figure 5 shows that patients in all stages die fairly quickly from hepatic failure compared to the other causes.

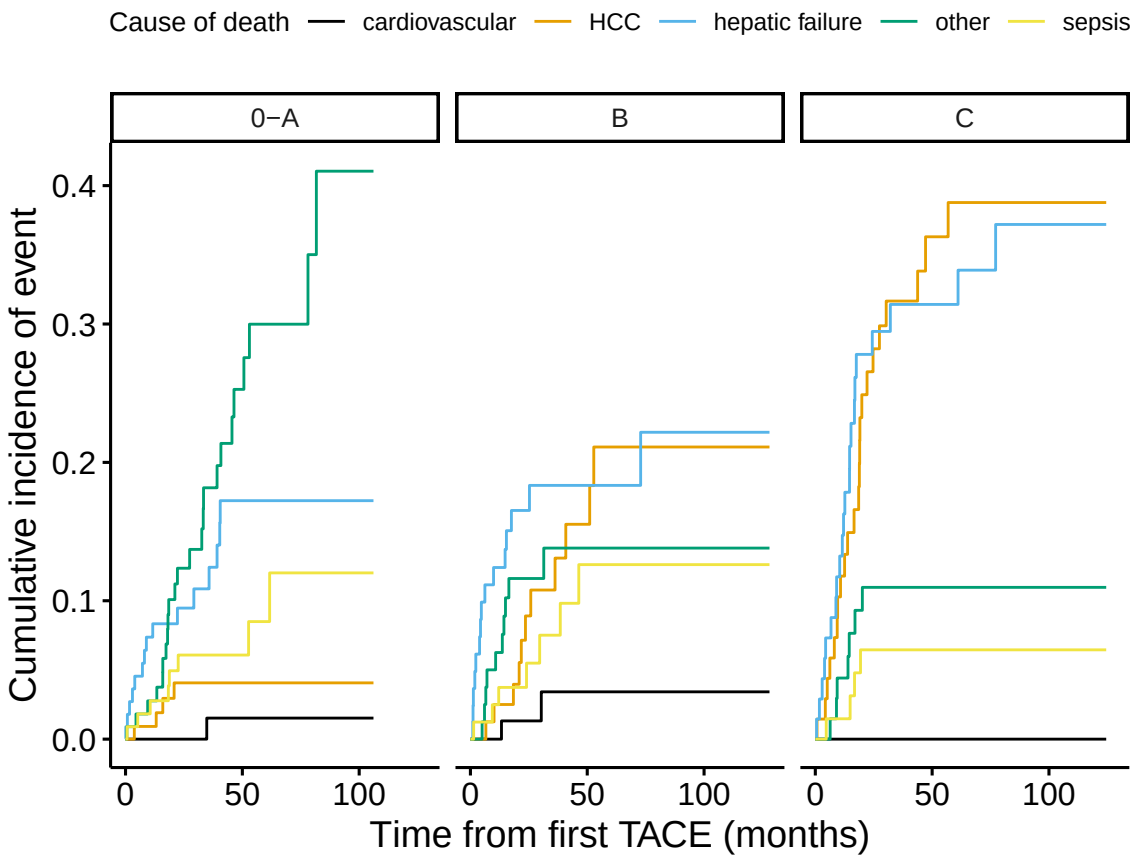

Figure 5: Cumulative incidence curves of competing risks of death since the first TACE treatment for the different BCLC stages

Table 5: Estimated cumulative incidence of causes of death by BCLC stages at 12 months, 24 months and 60 months. (HCC: hepatocellular carcinoma; Other: contains all causes of death that are not cardiovascular, HCC, hepatic failure or sepsis)

BCLC	**Cause of death**	**t = 12**	**t = 24**	**t = 60**
0-A	cardiovascular	0.0%	0.0%	1.5%
0-A	HCC	0.9%	4.1%	4.1%
0-A	hepatic failure	8.3%	9.5%	17.2%
0-A	other	2.8%	12.4%	30.0%
0-A	sepsis	2.8%	6.1%	8.5%
B	cardiovascular	0.0%	1.3%	3.4%
B	HCC	2.5%	8.9%	21.1%
B	hepatic failure	12.4%	16.5%	18.3%
B	other	6.3%	11.6%	13.8%
B	sepsis	2.5%	3.7%	12.6%
C	cardiovascular	0.0%	0.0%	0.0%
C	HCC	11.8%	26.5%	38.8%
C	hepatic failure	14.7%	27.8%	31.4%
C	other	4.4%	11.0%	11.0%
C	sepsis	1.5%	6.5%	6.5%

Table 5 shows values of the estimated cumulative incidence curves of causes of death by BCLC stages for 12 months, 24 months and 60 months after first TACE treatment. The probability for patients in stages 0-A to die from other causes increased from 12.4% at 24 months to 30.0% at 60 months. Contrarily, the probability for patients to die from other causes in stages B and C is roughly the same at 24 months compared to 60 months. Similarly, patients in stages 0-A have roughly the same risk to die from HCC at 24 months compared to 60 months (4.1%).

Overall Survival Hazard Ratios of Different Stages After Adjustment For Possible Confounders

I fitted a Cox-regression on the unadjusted OS for the BCLC stages. Table 6 shows that there is a tendency for increased OS for patients in stages 0-A when compared to patients in stage B (hazard ratio 0.74), however there is no evidence for increased OS ($p = 0.14$). Contrarily, there is strong evidence for decreased OS for patients in stage C when compared to patients in stage B as the risk to die at any point in time is almost doubled.

Table 6: Hazard ratios with 95%-confidence intervals and p-values estimated from the Cox-regression on the OS by BCLC stage without adjusting for possible confounders

	Hazard Ratio	95%-confidence interval	p-value
BCLC_EVENT0-A	0.74	from 0.49 to 1.10	0.14
BCLC_EVENTC	1.90	from 1.28 to 2.82	0.001

Next, I fitted a Cox-regression that adjusts for the possible confounders. Figure 6 summarizes the results from the regression. First, the results from adjusted Cox-regression for the effects of the BCLC stage are consistent with the results of the unadjusted Cox-regression. Second, our data suggests that there is no evidence for an association between the OS and sex, underlying disease, age and time to first TACE. Third, there was strong evidence for a difference in OS for patients that were treated with RFA/MWA before TACE when compared to patients who were not. Lastly, as expected the association between transplantation and OS is very strong and clear as the risk of dying is reduced by 74.0% (95%-confidence interval from 51.0% to 86.0% reduction) with transplantation.

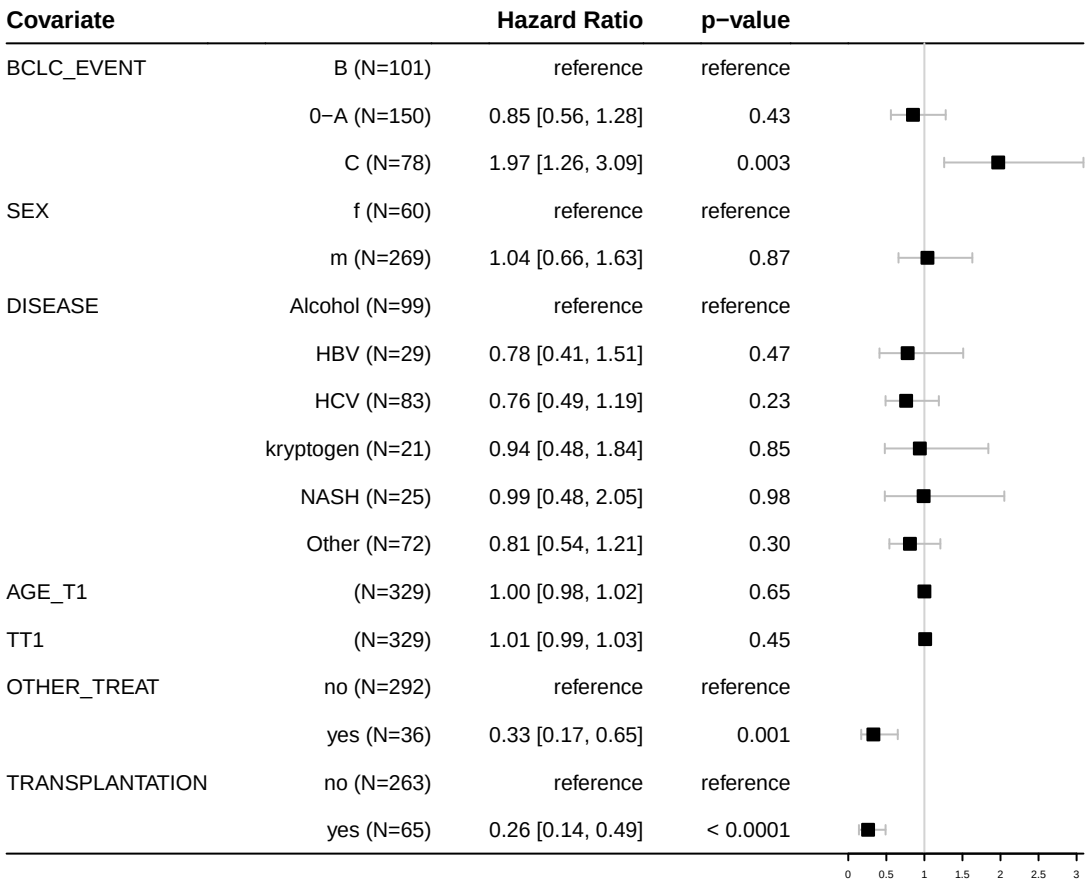

Figure 6: Hazard ratios with 95%-confidence intervals and p-values estimated from the Cox-regression on the OS by BCLC stage after adjusting for possible confounders (DISEASE: underlying condition; AGE_T1: age at first TACE treatment; TT1: time from diagnosis until first TACE treatment; OTHER_TREAT: patient received RFA and/or MWA treatment; TRANSPLANTATION: patient had a liver transplantation (time-varying covariate); HBV: hepatitis B virus; HCV: hepatitis C virus; NASH: non-alcoholic steatohepatitis)

As with all statistical models Cox-regression comes with its own set of assumptions. The main assumption is the proportional hazards assumption which states that the hazard ratio is constant over time. I checked the proportional hazards assumption with a visualization of the scaled Schoenfeld residuals. Figure 7 shows how the Betas change over time. In this case there is no evidence against the proportional hazards assumption as the smoothed line is approximately horizontal.

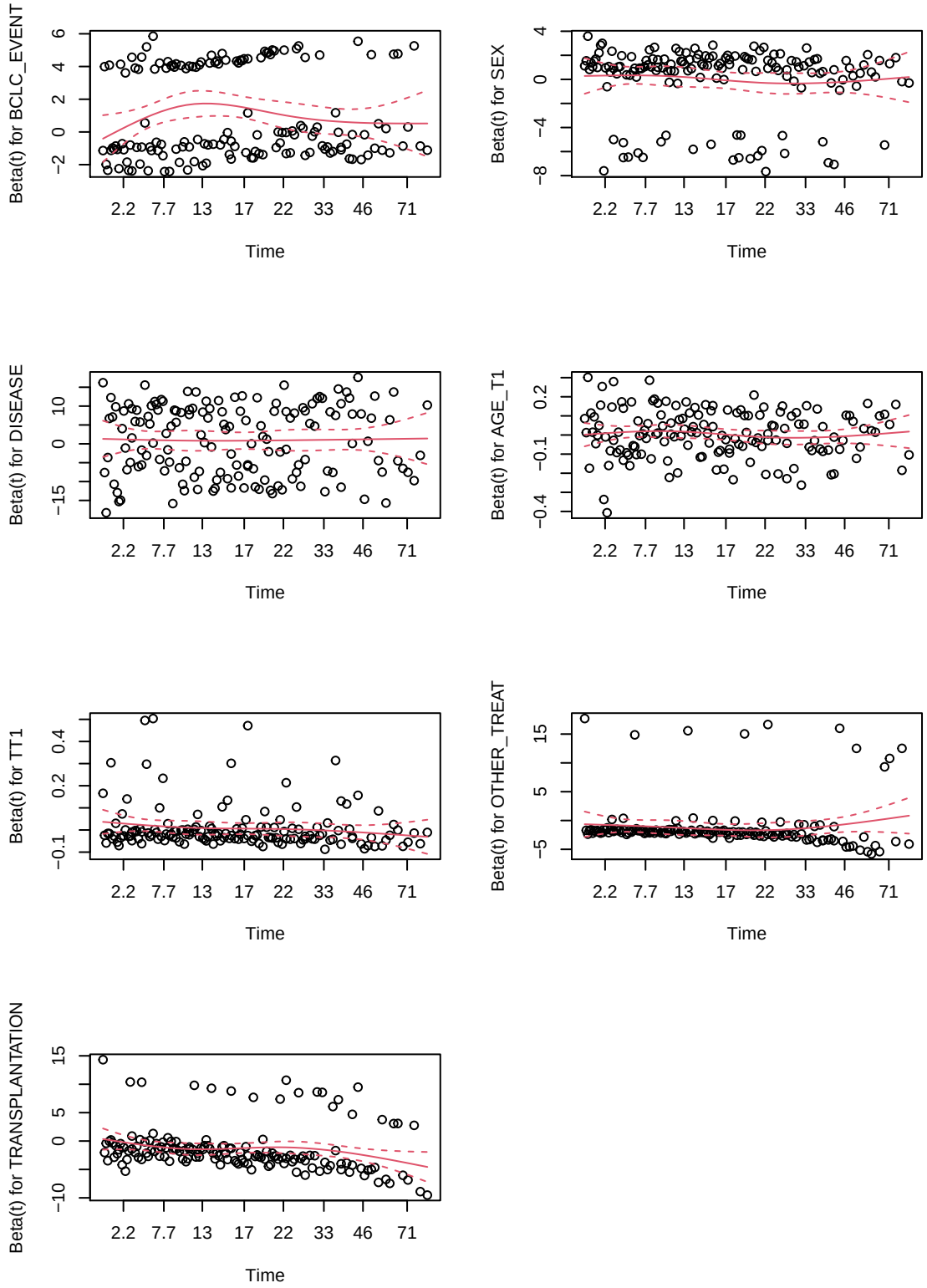

Figure 7: Analysis of the scaled Schoenfeld residuals to check for the proportional hazards assumption

6 Conclusion

In this study we looked at a very special patient population as only patients treated with TACE were included in the data set. We can assume that the patients in stages 0-A and C do not represent a random sample of all patients in said stages, because the patients in those stages were chosen for TACE according to some disease-specific properties. Thus, patients in stages 0-A in this study are probably more ill than the average patient in this stage. Vice versa patients in stage C are probably less ill than the average patient in this stage. Therefore, we have to consider that these assumptions influence all our interpretations.

In this report we compared our results to the guideline papers: Llovet et al. (2012), Llovet et al. (2018) and Forner et al. (2018). The more recent studies from Llovet et al. (2018) and Forner et al. (2018) show that there is a tendency for increased OS for patients in stage B ($>$ 10 months) and C ($>$ 1 month) compared to the results from Llovet et al. (2012). However, said studies did not provide confidence intervals of their results. Therefore, we mainly used Llovet et al. (2012) as it allowed us to calculate 95%-confidence intervals for the difference in median OS.

First, the results do not show that TACE improves the OS for patients in stages 0-A. The observed median OS was 47.8 months whereas the guideline papers reported median OS larger than 60 months. Based on these results we would not recommended patients of stages 0-A TACE treatments instead of the usual first-line treatment. However, this conclusion is only valid if we assume that the stage 0-A patients in our analysis are a representative sample from all stage 0-A patients. As we argued above, this might not be the case. Contrarily, patients in stage B had an improved median OS (20 months) compared to the results from Llovet et al. (2012). However, the 95%-confidence interval (from -8.5 to 49.5 months) did not indicate whether there was evidence for a difference in median OS. Patients in stage C had an improved median OS (8.4 months) compared to the results from Llovet et al. (2012). Additionally, the 95%-confidence interval (from 4.6 to 16.8 months) indicates that there is evidence for a difference in median OS. Therefore, we would suggest that patients of stages B and C would benefit from prolonged and earlier treatment with TACE.

The results from the pairwise log-rank test show that there is very strong evidence that the PFS of patients in stage C differs from patients in the other stages ($p < 0.0001$, $p < 0.0001$). Contrarily, the results do not show evidence for different PFS between patients in stages 0 and A and B ($p = 0.19$).

The results from the competing risks analysis show that patients in all BCLC stages die quickly from hepatic failure. Further, patients in stages B and C mostly die from causes related to the liver (HCC and hepatic failure) whereas patients in stages 0-A mostly die from other causes unrelated to the liver.

The unadjusted and adjusted Cox-regression show that the estimated effects of the BCLC stages are consistent. Our data indicates that the risk of dying is almost doubled for patients in BCLC stage C compared to patients in BCLC stage B. Further, the risk of dying is reduced by 74.0% when the patient underwent a liver transplantation.

References

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* **57** 289–300.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* **34** 187–202.

Forner, A., Reig, M. and Bruix, J. (2018). Hepatocellular carcinoma. *The Lancet* **391** 1301 – 1314.
URL http://www.sciencedirect.com/science/article/pii/S0140673618300102

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association* **53** 457–481.

Lévesque, L. E., Hanley, J. A., Kezouh, A. and Suissa, S. (2010). Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *Bmj* **340** b5087.

Llovet, J. M., Lencioni, R., Di Bisceglie, A. M., Gaile, P. R., Dufour, J. F., Greten, T. F., Raymond, E., Roskams, T., De Baere, T., Ducreux, M. et al. (2012). Easl-eortc clinical practice guidelines: management of hepatocellular carcinoma. *Journal of hepatology* **56** 908–943.

Llovet, J. M., Lencioni, R., Di Bisceglie, A. M., Gaile, P. R., Dufour, J. F., Greten, T. F., Raymond, E., Roskams, T., De Baere, T., Ducreux, M. et al. (2018). Easl clinical practice guidelines: management of hepatocellular carcinoma. *Journal of hepatology* **69** 182–236.

Peto, R. and Peto, J. (1972). Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society: Series A (General)* **135** 185–198.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL https://www.R-project.org/

Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C. and Vandenbroucke, J. P. (2007). The strengthening the reporting of observational studies in epidemiology (strobe) statement: guidelines for reporting observational studies. *Annals of internal medicine* **147** 573–577.

Zgraggen, A. (2018). *The Establishment of a Cohort and Quality Documentation for Patients with Hepatocellular Carcinoma at the University Hospital of Zurich*. Master's thesis, Universtiy of Zurich.

R version and packages used to generate this report:

R version: R version 4.0.3 (2020-10-10)

Base packages: grid, stats, graphics, grDevices, utils, datasets, methods, base

Other packages: ggthemes 4.2.0, forestplot 1.10, checkmate 2.0.0, magrittr 1.5, kableExtra 1.2.1, biostatUZH 1.8.0, gridExtra 2.3, survminer 0.4.8, ggpubr 0.4.0, xtable 1.8-4, tableone 0.12.0, cmprsk 2.2-10, survival 3.2-7, labelled 2.7.0, forcats 0.5.0, stringr 1.4.0, dplyr 1.0.2, purrr 0.3.4, readr 1.4.0, tidyr 1.1.2, tibble 3.0.4, ggplot2 3.3.2, tidyverse 1.3.0, readxl 1.3.1, knitr 1.30

This document was generated on Januar 12, 2021 at 16:07.

7 Appendix

```
##############################################
####### SETUP FOR THE REPORT FILE
##############################################
require(knitr)
opts_chunk$set(
fig.path = "plots/", echo = FALSE, message = FALSE,
results = "hide", fig.height = 5, fig.width = 6
)

##############################################
####### HELPER FUNCTIONS FOR CALCULATIONS AND
####### PLOTS
##############################################
# square-and-add method to construct 95%-CI for median difference
square_add <- function (m1, l1, u1, m2, l2, u2) {
  d <- m1 - m2
  l <- d - sqrt((m1 - l1)^2 + (u2 - m2)^2)
  u <- d + sqrt((u1 - m1)^2 + (m2 - l2)^2)
  return (c(lower = l, median = d, upper = u))
}

##############################################
####### DATA PREAPARATION
##############################################
# load packages
require(readxl) # package used for the reading of the excel file
require(tidyverse) # package used for loading tidy universe packages e.g. ggplot2, dplyr
require(labelled) # package used to add labels to data
require(survival) # package for the survival analysis calculations
require(cmprsk) # package used for the competing risk analysis
require(tableone) # package used for creating the baseline characteristics tables
require(xtable) # package used for tables
require(survminer) # package used for survival plots
require(gridExtra) # package used for aligning plots
require(biostatUZH) # package used to create tables
require(kableExtra) # package used to create tables
require(forestplot) # package for visualization of cox-regression
require(ggthemes) # package for themes and colorblind visualization

# additional settings
options(width = 85, digits = 3, show.signif.stars = FALSE)

# data path
data_path <- "../data/"

# path to the original excel file with the data
df_path <- paste0(data_path, "Datensatz TACE.xlsx")
 # sheet name where the data in the excel is located
sheet_name <- "Tabelle2"
# read the raw data from the excel file
raw_df <- read_excel(path = df_path, sheet = sheet_name, skip = 2)

# prepare data, only choose these variables from the raw data
df <- raw_df %>% select(SEX, BCLC, DISEASE, CAUSE, PROGRESS, DoBIRTH, DoHCC,
                        TACE1, TACE2, TACE3, TACE4, TACE5, TACE6, TACE7,
```

```
                          DLS, DEAD, DoDEATH, DoP, TRANSPLANTATION, RESECTION, METASTASIS,
                        ANGIOINVASION, BIOPSY, `NODULES #`, NODULEmax, RFA, MWA)

# assign arbitrary ID to each patient
df$ID <- 1:nrow(df)

# most common underlying causes in the data
diseases <- c("Alcohol", "HCV", "HBV", "kryptogen", "NASH")

# assign number of TACE per patient
df$NoT <- apply(df, 1, FUN = function (row) sum(!is.na(row[c("TACE1", "TACE2", "TACE3",
                                                             "TACE4", "TACE5", "TACE6",
                                                             "TACE7")])))

# converison of the transplantation variable to an actual date
df$TRANSPLANTATION_DATE <- as.POSIXct(ifelse(grepl("yes", df$TRANSPLANTATION),
                                             gsub(".*[ ]|[ ].*", "", df$TRANSPLANTATION), NA),
                                      format = "%d.%m.%Y", tz = "UTC")

# number to get time measure of unit month
time_divider <- 30.4

# prepare variables in data frame for analysis
df <- df %>% mutate(LAST_TIME = case_when(DEAD == "yes" ~ DoDEATH,
                                          TRUE ~ DLS),
                    EVENT = case_when(DEAD == "yes" ~ TRUE, TRUE ~ FALSE),
                    PROGRESS_EVENT = case_when((PROGRESS == "yes" | DEAD == "yes") ~ TRUE,
                                               TRUE ~ FALSE),
                    EVENT_TIME_DIAG = as.numeric(difftime(LAST_TIME, DoHCC, units = "days")/
                                                   time_divider),
                    EVENT_TIME_TACE = as.numeric(difftime(LAST_TIME, TACE1, units = "days")/
                                                   time_divider),
                    PROGRESS_TIME_DIAG = case_when(PROGRESS == "yes" ~
                                                     as.numeric(difftime(DoP,DoHCC, units = "days")
                                                                )/time_divider,
                                                   (PROGRESS == "no" & DEAD == "yes") ~
                                                     as.numeric(difftime(DoDEATH, DoHCC, units = "days")
                                                                )/time_divider,
                                                   TRUE ~ as.numeric(difftime(LAST_TIME, DoHCC,
                                                                              units = "days")
                                                                     )/time_divider),
                    PROGRESS_TIME_TACE = case_when(PROGRESS == "yes" ~
                                                     as.numeric(difftime(DoP, TACE1,
                                                                         units = "days")
                                                                )/time_divider,
                                                   (PROGRESS == "no" & DEAD == "yes") ~
                                                     as.numeric(difftime(DoDEATH, TACE1,
                                                                         units = "days")
                                                                )/time_divider,
                                                   TRUE ~ as.numeric(difftime(LAST_TIME, TACE1,
                                                                              units = "days")
                                                                     )/time_divider),
                    BCLC_EVENT = ifelse(BCLC == "0" | BCLC == "A", "0-A", BCLC),
                    NoT_GUIDE = case_when(NoT <= 2 ~ "Guideline", TRUE ~ "Extended"),
                    AGE_DIAGNOSIS = as.numeric(floor(difftime(DoHCC, DoBIRTH,
                                                              units = "days")/365)),
                    AGE_T1 = as.numeric(floor(difftime(TACE1, DoBIRTH, units = "days")/365)),
```

```
                        TT1 = as.numeric(difftime(TACE1, DoHCC, units = "days"))/time_divider,
                        TT2 = as.numeric(difftime(TACE2, TACE1, units = "days"))/time_divider,
                        TRANSPLANTATION = as.factor(case_when(grepl("yes", TRANSPLANTATION) ~ "yes",
                                                              TRUE ~ TRANSPLANTATION)),
                        TRANSPLANTATION_TIME_TACE = as.numeric(difftime(TRANSPLANTATION_DATE, TACE1,
                                                                        units = "days")/time_divider),
                        DISEASE = as.factor(case_when(DISEASE %in% diseases ~ DISEASE,
                                                      TRUE ~ "Other")),
                        RESECTION = as.factor(RESECTION),
                        METASTASIS = as.factor(METASTASIS),
                        ANGIOINVASION = as.factor(ANGIOINVASION),
                        BIOPSY = as.factor(case_when(grepl("yes", BIOPSY) ~ "yes", TRUE ~ BIOPSY)),
                        CC = case_when(EVENT ~ CAUSE, TRUE ~ "No event"),
                        TIME_START_A  = 0,
                        TIME_END_A = case_when(TRANSPLANTATION == "yes" ~ TRANSPLANTATION_TIME_TACE,
                                               TRUE ~ EVENT_TIME_TACE),
                        TIME_START_B = case_when(TRANSPLANTATION == "yes" ~ TRANSPLANTATION_TIME_TACE,
                                                 TRUE ~ NA_real_),
                        TIME_END_B = case_when(TRANSPLANTATION == "yes" ~ EVENT_TIME_TACE,
                                               TRUE ~ NA_real_),) %>%
  filter(LAST_TIME > TACE1, case_when(TRANSPLANTATION == "yes" ~ TRANSPLANTATION_DATE > DoHCC,
                                      TRUE ~ TRUE)) # filter out invalid observations

# change all character variables to factors
df[sapply(df, is.character)] <- lapply(df[sapply(df, is.character)],
                                       as.factor)

# relevel variable CC (competing risk)
df$CC <- factor(df$CC, levels = c("No event", "cardiovascular", "HCC",
                                  "hepatic failure", "other", "sepsis"))

# create new variable based on if patients had RFA/MWA
df <- df %>% rowwise() %>% mutate(OTHER_TREAT = as.factor(case_when(all(c(is.na(RFA), is.na(MWA))) ~
                                                                      NA_character_,
                                                                  any(c(grepl("yes", RFA),
                                                                        grepl("yes", MWA))) ~ "yes",
                                                                  TRUE ~ "no")))

# select only those variables in data frame
df <- df %>% select(SEX, BCLC, BCLC_EVENT, DISEASE, CAUSE, PROGRESS_EVENT,
                    PROGRESS_TIME_DIAG, PROGRESS_TIME_TACE, EVENT, EVENT_TIME_DIAG, EVENT_TIME_TACE,
                        AGE_DIAGNOSIS, AGE_T1, TT1, TT2, NoT, NoT_GUIDE, TRANSPLANTATION, RESECTION,
                    METASTASIS, ANGIOINVASION, BIOPSY, `NODULES #`, NODULEmax, OTHER_TREAT, CC,
                    TIME_START_A, TIME_END_A, TIME_START_B, TIME_END_B, ID, LAST_TIME)

# create labels for data frame
var_label(df) <- list(
  SEX = "Sex",
  BCLC = "BCLC stage",
  BCLC_EVENT = "",
  DISEASE = "Underlying cause for HCC",
  CAUSE = "Cause of death",
  PROGRESS_EVENT = "",
  PROGRESS_TIME_DIAG = "",
  PROGRESS_TIME_TACE = "",
  EVENT = "",
  EVENT_TIME_DIAG = "",
```

```
  EVENT_TIME_TACE = "",
  AGE_DIAGNOSIS = "Age at diagnosis",
  AGE_T1 = "Age at first TACE",
  TT1 = "Time from diagnosis until first TACE (in months)",
  TT2 = "Time from first TACE to second TACE (in months)",
  NoT = "Number of TACE treatments received",
  NoT_GUIDE = "Patient has received TACE treatment according to guidelines (2 or less)",
  TRANSPLANTATION = "Liver transplantation after TACE",
  RESECTION = "Liver resection before TACE",
  METASTASIS = "Metastases before TACE",
  ANGIOINVASION = "Angioinvasion before TACE",
  BIOPSY = "Biopsy before TACE",
  `NODULES #` = "Number of nodules/lesions before TACE",
  NODULEmax = "Size of the largest nodule/lesion before TACE",
  OTHER_TREAT = "Patient has received other treatment additional to TACE",
  CC = "Competing risk event",
  ID = "Arbitrary Patient ID",
  TIME_START_A = "",
  TIME_END_A = "",
  TIME_START_B = "",
  TIME_END_B = ""
)

##############################################
####### DESCRIPTIVE STATISTICS & MODELLING
##############################################
# sex distribution of the patients
g_distr <- ggplot(df, aes(x = SEX, y = ..count../sum(..count..))) +
  geom_bar() + xlab("Sex of the patient") + ylab("Relative frequency") +
  scale_x_discrete(labels = c("Female", "Male")) +
  scale_y_continuous(labels = scales::percent, limits = c(0, 1)) +
  ggtitle("Gender distribution of the patients") +
  geom_text(stat = "count", aes(label = paste0(round(100*..count../sum(..count..),
                                                digits = 1), "%")), vjust = -1) +
  theme_bw()

# BCLC stage distribution of the patients
s_distr <- ggplot(df, aes(x = BCLC, y = ..count../sum(..count..))) +
  geom_bar() + xlab("BCLC stage of the patient") + ylab("Relative frequency") +
  scale_y_continuous(labels = scales::percent, limits = c(0, 0.4)) +
  ggtitle("BCLC stage distribution of the patients") +
  geom_text(stat = "count", aes(label = paste0(round(100*..count../sum(..count..),
                                                digits = 1), "%")),
            vjust = -1, position = "stack") +
  theme_bw()

# align both plots
grid_base_line <- grid.arrange(g_distr, s_distr,
                               nrow = 2)
grid_base_line
# create baseline characteristics table
tab_base <- CreateTableOne(data = df, strata = "BCLC",
                           vars = c("SEX", "AGE_DIAGNOSIS", "DISEASE"),
                           test = FALSE)
tab_base <- print(tab_base, printToggle = FALSE, noSpaces = TRUE,
                  varLabels = TRUE, floating = FALSE)
kable(tab_base, format = "latex",
```

```
      caption = "Patient characteristics at diagnosis by BCLC stages (SD: standard deviation;
      HCC: hepatocellular carcinoma; HBV: hepatitis B virus; HCV: hepatitis C virus;
      NASH: non-alcoholic steatohepatitis) \\label{tab:BASE_CHAR}",
      booktabs = TRUE) %>%
  kable_styling(latex_options = c("scale_down", "HOLD_position")) %>%
  row_spec(0, bold = TRUE)

# filter out patients with stage D
kpm <- df %>% filter(BCLC_EVENT != "D")

# relevel BCLC stages to have stage B as reference level
kpm$BCLC_EVENT <- factor(kpm$BCLC_EVENT, levels = sort(unique(kpm$BCLC_EVENT)))

# create patient characteristics table
tab_pt <- CreateTableOne(data = kpm, strata = "BCLC_EVENT",
                         vars = c("NoT", "NoT_GUIDE",
                                  "AGE_T1", "TT1", "TT2",
                                  "TRANSPLANTATION", "RESECTION",
                                  "METASTASIS", "ANGIOINVASION",
                                  "BIOPSY", "CC", "NODULES #",
                                  "NODULEmax", "OTHER_TREAT"),
                         test = FALSE)
tab_pt <- print(tab_pt, printToggle = FALSE, noSpaces = TRUE,
                varLabels = TRUE, floating = FALSE)
kable(tab_pt, format = "latex",
      caption = "Additional patient characteristics by BCLC stages after omitting the sole patient with
      BCLC stage D and grouping patients of stages 0 and A as one group (SD: standard deviation;
      HCC: hepatocellular carcinoma; HBV: hepatitis B virus; HCV: hepatitis C virus;
      NASH: non-alcoholic steatohepatitis; Patient received other treatment additional to TACE: patient
      received RFA and/or MWA) \\label{tab:PAT_CHAR}",
      booktabs = TRUE) %>%
  kable_styling(latex_options = c("scale_down", "HOLD_position")) %>%
  row_spec(0, bold = TRUE)
# calculate Kaplan-Meier estimates OS
kpm_fit_OS <- with(kpm, survfit(Surv(time = EVENT_TIME_DIAG, event = EVENT) ~ BCLC_EVENT))

# plot Kaplan-Meier curves
kpm_OS_plot <- ggsurvplot(kpm_fit_OS, data = kpm, surv.median.line = "hv", legend.title = "BCLC stage",
                 legend.labs = c("0-A (early)",
                                 "B (intermediate)",
                                 "C (advanced)"), conf.int = TRUE,
                 xlab = "Time from diagnosis (months)", ylab = "Overall survival (%)")$plot

kpm_OS_plot + scale_y_continuous(labels = scales::percent, limits = c(0, 1)) +
  scale_color_colorblind() + scale_fill_colorblind()

# calculate CI for difference in median OS
m <- summary(kpm_fit_OS)$table[, "median"]
l <- summary(kpm_fit_OS)$table[, "0.95LCL"]
u <- summary(kpm_fit_OS)$table[, "0.95UCL"]

m_B <- 20; l_B <- 14; u_B <- 45;
m_C <- 11; l_C <- 6; u_C <- 14;

# no evidence for a significant median difference in OS
diff_B <- square_add(m1 = m[2], l1 = l[2], u1 = u[2],
                     m2 = m_B, l2 = l_B, u2 = u_B)
```

```
# evidence for a significant median difference in OS
diff_C <- square_add(m1 = m[3], l1 = l[3], u1 = u[3],
                     m2 = m_C, l2 = l_C, u2 = u_C)

# table to compare observed data to other studies
table_text_A <- data.frame(X1 = c("Observed data", "Llovet et al. (2012)", "Llovet et al. (2018)",
                                  "Difference Observed and Llovet et al. (2012)"),
                           X2 = c(round(summary(kpm_fit_OS)$table[1, "median"], digits = 1),
                                  "$\\geq$60", "$\\geq$60", "-"),
                           X3 = c(paste("from", round(summary(kpm_fit_OS)$table[1, "0.95LCL"],
                                                      digits = 1),
                                        "to", round(summary(kpm_fit_OS)$table[1, "0.95UCL"],
                                                    digits = 1)),
                                  "-", "-", "-"))

table_text_A5 <- data.frame(X1 = c("Observed data", "Llovet et al. (2012)"),
                           X2 = c(paste0(100*round(summary(kpm_fit_OS, time = 60)$surv[1],
                                                   digits = 2), "\\%"), "-"),
                           X3 = c(paste("from", paste0(100*round(summary(kpm_fit_OS,
                                                                         time = 60)$lower[1],
                                                                 digits = 2), "\\%"), "to",
                                        paste0(100*round(summary(kpm_fit_OS, time = 60)$upper[1],
                                                         digits = 2), "\\%")),
                                  "from 40\\% to 70\\%"))

table_text_B <- data.frame(X1 = c("Observed data", "Llovet et al. (2012)", "Llovet et al. (2018)",
                                  "Difference Observed and Llovet et al. (2012)"),
                           X2 = c(round(summary(kpm_fit_OS)$table[2, "median"], digits = 1),
                                  20, "$\\geq$30",
                                  round(diff_B[2], digits = 1)),
                           X3 = c(paste("from", round(summary(kpm_fit_OS)$table[2, "0.95LCL"],
                                                      digits = 1), "to",
                                        round(summary(kpm_fit_OS)$table[2, "0.95UCL"], digits = 1)),
                                  "from 14 to 45", "-", paste("from", round(diff_B[1], digits = 1),
                                                              "to",
                                                              round(diff_B[3], digits = 1))))

table_text_C <- data.frame(X1 = c("Observed data", "Llovet et al. (2012)", "Llovet et al. (2018)",
                                  "Forner et al. (2018)",
                                  "Difference Observed and Llovet et al. (2012)"),
                           X2 = c(round(summary(kpm_fit_OS)$table[3, "median"], digits = 1), 11,
                                  "$\\geq$10", "$\\geq$12",
                                  round(diff_C[2], digits = 1)),
                           X3 = c(paste("from", round(summary(kpm_fit_OS)$table[3, "0.95LCL"],
                                                      digits = 1), "to",
                                        round(summary(kpm_fit_OS)$table[3, "0.95UCL"], digits = 1)),
                                  "from 6 to 14", "-", "-",
                                  paste("from", round(diff_C[1], digits = 1), "to",
                                        round(diff_C[3], digits = 1))))

table_full <- rbind(table_text_A, table_text_B, table_text_C,
                    data.frame(X1 = "Study", X2 = "5-year OS", X3 = "95\\%-confidence interval"),
                    table_text_A5)
colnames(table_full) <- c("Study", "Median OS", "95\\%-confidence interval")
kable(table_full, format = "latex", escape = FALSE, booktabs = TRUE,
      align = c("l", "r", "r"),
      caption = "Observed median overall survival time (in months) compared to available data from
```

```
      other studies for all BCLC stages and the observed 5-year overall survival compared to available
      data from
      other studies for BCLC stages 0-A \\label{tab:KPM_OS}") %>%
  kable_styling(latex_options = c("scale_down", "HOLD_position")) %>% pack_rows("Stages 0-A", 1, 4,
                                                                                bold = FALSE) %>%
  pack_rows("Stage B", 5, 8, bold = FALSE) %>% pack_rows("Stage C", 9, 13, bold = FALSE) %>%
  row_spec(0, bold = TRUE) %>%
  row_spec(14, bold = TRUE, extra_latex_after = "\\hline") %>%
  pack_rows("Stages 0-A", 15, 16, bold = FALSE)
# log-rank test for OS
pair_log_rank <- pairwise_survdiff(Surv(time = EVENT_TIME_DIAG,
                                        event = EVENT) ~ BCLC_EVENT, data = kpm, p.adjust.method = "BH")
# calculate Kaplan-Meier estimates PFS
kpm_fit_PFS <- with(kpm, survfit(Surv(time = PROGRESS_TIME_DIAG, event = PROGRESS_EVENT) ~ BCLC_EVENT))
# plot Kaplan-Meier curves
kpm_PFS_plot <- ggsurvplot(kpm_fit_PFS, data = kpm, surv.median.line = "hv",
                           legend.title = "BCLC stage",
                 legend.labs = c("0-A (early)",
                                 "B (intermediate)",
                                 "C (advanced)"), conf.int = TRUE,
                 xlab = "Time from diagnosis (months)", ylab = "Progression-free survival (%)")$plot
kpm_PFS_plot + scale_y_continuous(labels = scales::percent, limits = c(0, 1)) +
  scale_color_colorblind() + scale_fill_colorblind()
# log-rank test for PFS
pair_log_rank_PFS <- pairwise_survdiff(Surv(time = PROGRESS_TIME_DIAG, event = PROGRESS_EVENT) ~
                                       BCLC_EVENT, data = kpm, p.adjust.method = "BH")
# initialize data
kpm$BCLC_EVENT <- relevel(kpm$BCLC_EVENT, ref = "B")

# competing risks plot
comp <- cuminc(ftime = kpm$EVENT_TIME_TACE, fstatus = factor(as.numeric(kpm$CC) - 1),
               group = kpm$BCLC_EVENT)

ggcompetingrisks(comp, xlab = "Time from first TACE (months)",
                               ylab = "Cumulative incidence of event",
                               title = "",
                 conf.int = FALSE) +
  scale_color_colorblind(name = "Cause of death", labels = levels(kpm$CAUSE)) +
  theme(legend.text = element_text(size = 9))
# calculate CI for competing risk estimates
comp_est <- timepoints(comp, times = c(c(1, 2, 5)*12))$est %>%
  as.data.frame() %>%
  mutate(BCLC_EVENT = sub(" .*", "", rownames(.)),
         EVENT = rep(levels(kpm$CAUSE), each = 3)) %>%
  transmute(BCLC = BCLC_EVENT,
         CAUSE = EVENT,
         T12 = formatPercent(`12`),
         T24 = formatPercent(`24`),
         T60 = formatPercent(`60`)) %>%
  arrange(BCLC)

kable(comp_est, format = "latex", col.names = c("BCLC", "Cause of death",
                                                "t = 12", "t = 24", "t = 60"),
      booktabs = TRUE, escape = TRUE,
      caption = "Estimated cumulative incidence of causes of death by BCLC stages at 12 months,
      24 months and 60 months. (HCC: hepatocellular carcinoma; Other: contains all causes of death
      that are not cardiovascular, HCC, hepatic failure or sepsis) \\label{tab:COMP}") %>%
```

```
  kable_styling(latex_options = "HOLD_position") %>%
  row_spec(0, bold = TRUE)
# unadjusted Cox-regression on the OS
fit_cox_bas <- coxph(Surv(EVENT_TIME_TACE, EVENT) ~ BCLC_EVENT, data = kpm)
tableRegression(fit_cox_bas, caption = "Hazard ratios with 95\\%-confidence intervals and
                $p$-values estimated from the Cox-regression on the OS by BCLC stage without adjusting
                for possible confounders",
                caption.placement = "top",
                label = "tab:REG_BASE",
                table.placement = "H")

# prepare time-varying data based on transplantation
df_long <- kpm %>%
  unite("A", TIME_START_A:TIME_END_A, na.rm = FALSE, remove = TRUE, sep = ";") %>%
  unite("B", TIME_START_B:TIME_END_B, na.rm = FALSE, remove = TRUE, sep = ";") %>%
  gather(TYPE, TIME, A:B) %>%
  separate(TIME, into = c("START", "END"), sep = ";", convert = TRUE) %>%
  filter(!is.na(START), !is.na(END))

ids <- df_long %>% filter(TRANSPLANTATION == "yes") %>% pull(ID) %>% unique()

df_long <- df_long %>% mutate(TRANSPLANTATION = case_when((ID %in% ids & TYPE == "A") ~ "no",
                                                          TRUE ~ as.character(TRANSPLANTATION)),
                              EVENT = case_when((ID %in% ids & TYPE == "A") ~ FALSE,
                                                TRUE ~ EVENT)) %>%
  select(-TYPE)

# time-varying cox-regression
fit_cox <- coxph(Surv(START, END, EVENT) ~ BCLC_EVENT + SEX + DISEASE + AGE_T1 + TT1 +
                   OTHER_TREAT + TRANSPLANTATION + cluster(ID), data = df_long)

est <- round(exp(coef(fit_cox)), digits = 2)
low <- round(exp(confint(fit_cox, level = 0.95))[, 1], digits = 2)
upp <- round(exp(confint(fit_cox, level = 0.95))[, 2], digits = 2)
pval <- formatPval(summary(fit_cox)$coefficients[, 6])
table_forest <- cbind(
  c("Covariate", "BCLC_EVENT", rep("", 2),
    "SEX", "", "DISEASE", rep("", 5), "AGE_T1", "TT1",
    "OTHER_TREAT", "", "TRANSPLANTATION", ""),
  c("", "B (N=101)", "0-A (N=150)", "C (N=78)", "f (N=60)", "m (N=269)", "Alcohol (N=99)", "HBV (N=29)",
    "HCV (N=83)", "kryptogen (N=21)", "NASH (N=25)", "Other (N=72)", "(N=329)",
    "(N=329)", "no (N=292)", "yes (N=36)", "no (N=263)", "yes (N=65)"),
  c("Hazard Ratio", "reference", paste(sprintf("%.2f", est[1:2]), formatCI(cbind(low[1:2], upp[1:2]))),
    "reference", paste(sprintf("%.2f", est[3]), formatCI(cbind(low[3], upp[3]))), "reference",
    paste(sprintf("%.2f", est[4:8]), formatCI(cbind(low[4:8], upp[4:8]))),
    paste(sprintf("%.2f", est[9]), formatCI(cbind(low[9], upp[9]))),
    paste(sprintf("%.2f", est[10]), formatCI(cbind(low[10], upp[10]))), "reference",
    paste(sprintf("%.2f", est[11]), formatCI(cbind(low[11], upp[11]))), "reference",
    paste(sprintf("%.2f", est[12]), formatCI(cbind(low[12], upp[12])))),
  c("p-value", "reference", pval[1:2], "reference", pval[3],
    "reference", pval[4:8], pval[9], pval[10], "reference", pval[11],
    "reference", pval[12])
)

forestplot(table_forest,
           mean = c(NA, 1, est[1:2], 1, est[3],
             1, est[4:8], est[9], est[10],
```

```
             1, est[11], 1, est[12]),
           lower = c(NA, NA, low[1:2], NA, low[3],
              NA, low[4:8], low[9], low[10],
              NA, low[11], NA, low[12]),
           upper = c(NA, NA, upp[1:2], NA, upp[3],
              NA, upp[4:8], upp[9], upp[10],
              NA, upp[11], NA, upp[12]),
           boxsize = 0.25,
           zero = 1,
           vertices = TRUE,
           is.summary = c(TRUE, rep(FALSE, 18)),
           hrzl_lines = list("2" = gpar(lty = 1),
                             "19" = gpar(lty = 1)),
           txt_gp = fpTxtGp(cex = 0.7))
# test proportional hazards assumption
cox_test <- cox.zph(fit_cox)

# visualize proportional hazards assumption test
par(mfrow = c(4, 2))
plot(cox_test, col = 2)
```