의학석사 학위논문

단일기관 코호트 연구와 다기관 코호트 연구의 간세포암 환자 치료별 생존률의 비교 및 해석

Comparison of Overall Survival by Treatment in Hepatocellular Cancer Patients between Single-Center Cohort Study and Multi-Center Cohort Study

울산대학교 대학원

의 학 과

김 예 림

단일기관 코호트 연구와 다기관 코호트 연구의 간세포암 환자 치료별 생존률의 비교 및 해석

지 도 교 수 심 주 현

이 논문을 의학석사 학위 논문으로 제출함

2024 년 2 월

울산대학교 대학원

의 학 과

김 예 림

김예림의 의학석사 학위 논문을 인준함

심사위원　　이 한 주　　인

심사위원　　심 주 현　　인

심사위원　　이 단 비　　인

울 산 대 학 교　대 학 원

2024 년 2 월

국문요약

연구배경: 현재까지 단일 기관 및 다기관 후향적 코호트 연구의 치료 경향 및 생존률 간 유의한 차이가 관찰되는지에 대해서는 직접적인 비교검증 연구가 없는 상황이다. 이에 본 연구에서는 간암등록자료를 사용한 전국 코호트와 대표성 있는 단일기관 코호트의 자료를 비교하여 치료별 생존률의 유의한 차이가 있는지 알아보고자 하였다.

연구 방법: 2008 년부터 2018 년 사이 대한간암학회 간암등록사업에 등록된 16,443 명의 간세포암 환자들과 서울아산병원 간세포암 레지스트리에 등록된 15,655 명의 간세포암 환자들을 후향적으로 분석하였다. 일차 평가 변수는 전체 생존률로 정의하여, 초치료로 간절제, 간이식, 색전술, 국소절제요법, 항암화학요법, 방사선 치료, 그리고 완화의료를 시행한 각 환자군의 생존률을 분석하였고, 전체 환자 중 각 병기에 따라 Barcelona Clinic Liver Cancer (BCLC) 가이드라인에서 권고되는 치료를 시행한 군을 분류하여 단일기관 및 다기관 코호트 사이 치료별 생존률의 유의한 차이 여부를 비교하였다.

연구결과: 단일기관 코호트에서 추적관찰 기간의 중앙값은 36.2 개월 (사분위수 9.7-66.9 개월), 다기관 코호트에서는 30.0 개월 (사분위수 6.1-60.0 개월)이었고, 전체 환자군을 비교하였을 때 단일기관 코호트의 생존률 중앙값이 73.6 개월로 다기관 코호트의 중앙값인 34.0 개월보다 전체 생존률이 유의하게 높았다 (P<0.001). 치료별 생존률을 다변량 콕스 분석을 통해 비교하였을 때 다기관 코호트에서 완치요법 및 비완치요법을 받은 환자들에서 조정 위험률(adjusted hazard ratio)이 각각 1.48 (95% 신뢰구간 1.39-1.59), 1.22 (95% 신뢰구간 1.17-1.27)로 사망 위험이 유의하게 높았으며 완화의료를 받은 군의 사망 위험이 조정 위험률 0.85 (95% 신뢰구간 0.79-0.91)로 유의하게 낮았다. BCLC 가이드라인에 따른 치료별 생존률을 비교하였을 때도 다기관 코호트에서 색전술과 국소절제요법을 받은 군이 각각 조정 위험률이 1.72 (95% 신뢰구간 1.48-2.00), 1.44 (95% 신뢰구간 1.08-1.92)로 유의하게 낮은 생존률을 보였다.

연구결론: 단일기관 코호트와 다기관 코호트에서 간세포암 환자의 치료별 생존률을 비교하였을 때 전체 생존률의 유의한 차이가 관찰되어 향후 단일기관 코호트 연구의 간세포암 생존률 결과 해석에 대한 주의가 필요하다.

Contents

List of tables and figures iii

Introduction 1

Patients and Method 2

Study design and patient selection 2

Variables 2

Treatment modalities 3

Outcomes 3

Statistical analysis 4

Results 4

Study population 4

Distribution of liver function by initial treatment 10

Survival outcomes 12

Subcohort analysis of patients treated according to BCLC guidelines 31

Subgroup analysis 37

Discussion 40

Conclusion 43

Abstract 49

List of tables and figures

TABLE 1. Baseline characteristics of the study populations* 6
TABLE 2. Baseline characteristics of the study populations (missing data classified as a category) 8
TABLE 3. Distribution of liver function by initial treatment* 11
TABLE 4. Cox regression analysis of factors associated with mortality in the entire cohorts* 19
TABLE 5. Cox regression analysis of factors associated with mortality in the entire cohort (missing data classified as a category) 21
TABLE 6. Cox regression analysis of risk of mortality by initial treatment in the entire cohorts* 24
TABLE 7. Cox regression analysis of risk of mortality by initial treatment in the entire cohorts (missing data classified as a category) 25
TABLE 8. Cox regression analysis of risk of mortality by initial treatment in BCLC-guided sub-cohorts* 35
TABLE 9. Cox regression analysis of risk of mortality by initial treatment in the BCLC-guided subcohorts (missing data classified as a category) 36
TABLE 10. Subgroup analysis* 38

Figure 1. Patient flowchart of the study population 5
Figure 2. Kaplan–Meier estimates of overall survival in the two cohorts 12
Figure 3A. Kaplan-Meier estimates of overall survival according to male sex 13
Figure 3B. Kaplan-Meier estimates of overall survival according to female sex 13
Figure 4A. Kaplan-Meier estimates of overall survival according to Child-Pugh class A 14
Figure 4B. Kaplan-Meier estimates of overall survival according to Child-Pugh class B 14
Figure 4C. Kaplan-Meier estimates of overall survival according to Child-Pugh class C 15
Figure 5A. Kaplan-Meier estimates of overall survival according to modified UICC stage I 15
Figure 5B. Kaplan-Meier estimates of overall survival according to modified UICC stage II 16
Figure 5C. Kaplan-Meier estimates of overall survival according to modified UICC stage III 16
Figure 5D. Kaplan-Meier estimates of overall survival according to modified UICC stage IVA 17
Figure 5E. Kaplan-Meier estimates of overall survival according to modified UICC stage IVB 17
Figure 6A. Kaplan-Meier estimates of overall survival following surgical resection 26
Figure 6B. Kaplan-Meier estimates of overall survival following liver transplantation 26

Figure 6C. Kaplan-Meier estimates of overall survival following local ablation therapy 27
Figure 6D. Kaplan–Meier estimates of overall survival following transarterial chemoembolization/transarterial radioembolization 27
Figure 6E. Kaplan-Meier estimates of overall survival following systemic therapy 28
Figure 6F. Kaplan-Meier estimates of overall survival following radiotherapy 28
Figure 7A. Kaplan-Meier estimates of overall survival following curative treatment 29
Figure 7B. Kaplan-Meier estimates of overall survival following non-curative treatment 29
Figure 7C. Kaplan-Meier estimates of overall survival following best supportive care 30
Figure 8A. Kaplan–Meier estimates of overall survival of patients who received surgical resection according to the treatment indications 32
Figure 8B. Kaplan–Meier estimates of overall survival of patients who received liver transplant according to the treatment indications 32
Figure 8C. Kaplan–Meier estimates of overall survival of patients who received local ablation therapy according to the treatment indications 33
Figure 8D. Kaplan–Meier estimates of overall survival of patients who received transarterial chemoembolization according to the treatment indications 33
Figure 8E. Kaplan–Meier estimates of overall survival of patients who received systemic therapy according to the treatment indications 34

Introduction

Of the various observational study designs, retrospective cohort studies allow relatively quick, cost-effective, and practicable analyses of the associations between multiple exposures and the corresponding outcomes.[1] These outcomes are established on the basis of existing data for a representative patient population under broad inclusion criteria, thereby providing more generalizable results. In the clinical setting, retrospective cohort studies are especially important in hepatocellular carcinoma (HCC), a disease with heterogeneous tumor characteristics and treatment options depending on staging and underlying liver function.[2-4] A consequence of this is that randomized controlled trials (RCT) are less likely to reveal the variable clinical course of HCC and may not reflect real-world treatment outcomes. This highlights the need for observational studies that can objectively portray overall survival in actual clinical practice.[5] However, retrospective cohort studies have several limitations, of which external validity and selection bias are considered of major concern.[6]

Retrospective cohort studies are often conducted at a multicenter level to overcome this problem,[7] but the rationale for this approach is mostly based on evidence acquired from RCTs.[8] Upon comparison of previous RCTs, single-center trials have shown larger intervention effects than multicenter trials,[9, 10] or, in other cases, positive results of single-center trials have been contradicted by subsequent multicenter trials.[8, 11, 12] However, the differences in outcome seen in RCTs have not yet been demonstrated in retrospective cohort studies; due to the differences between RCTs and retrospective cohort studies in study design and patient population,[13] it is unclear whether retrospective single-center cohort studies have the same drawbacks as single-center RCTs. If the results of retrospective single-center cohort studies of adequate sample size are comparable to those of multicenter cohort studies, researchers might be spared the time and effort of achieving uniformity of data between different institutions while obtaining a similar degree of external validity.

We thus hypothesized that a well-conducted single-center study of adequate sample size could potentially establish the survival outcomes of HCC despite its variable disease course and tumor features. The objective of the present study was to evaluate the external validity and generalizability of retrospective single-center cohort studies by comparing the overall survival outcomes of a nationwide multicenter cohort and a large single-center cohort.

Patients and Method

Study design and patient selection

We conducted a retrospective analysis of de-identified patients newly diagnosed with HCC using data from a nationwide multicenter cohort and a single-center cohort in South Korea between January 2008 and December 2018. Diagnosis of HCC was made histologically or radiologically according to the criteria of the American Association for the Study of Liver Disease (AASLD), European Association for the Study of the Liver (EASL), and the Korean Liver Cancer Association (KLCA).[2-4] The Korean Primary Liver Cancer Registry (KPLCR) was selected as the multicenter cohort, and the Asan Medical Center (AMC) HCC registry, developed using the well-established Research Electronic Data Capture (REDCap) Cloud platform at South Korea's largest cancer institute and hospital (https://eng.amc.seoul.kr), was selected as the single-center cohort.[14, 15] The KPLCR is a database containing approximately 15% of patients newly-diagnosed with HCC registered in the South Korean Central Cancer Registry, from which patients are randomly selected each year using the probability proportional to size method and stratification by region (54 hospitals, with a variety of levels of care).[16] Eligible patients were male and female patients aged 18 years or over, and patients for whom no information was available regarding number of tumors, treatment modality, and age at diagnosis were excluded. This study was approved by the Institutional Review Board of the Asan Medical Center (IRB no.:2022-1274), which waived a requirement for informed consent owing to the retrospective nature of the study.

Variables

Baseline characteristics of the study population included age, sex, body mass index, underlying hypertension or diabetes mellitus, and presence of viral hepatitis, which was defined as any of the following: positive hepatitis B surface antigen or hepatitis C antibody, positive viral titer, or previous history of antiviral therapy. Baseline liver function was assessed by Child Pugh score and Model for End-stage Liver Disease (MELD) score. Tumors were staged according to the Barcelona Clinic Liver Cancer (BCLC) strategy, and the modified Union for International Cancer Control (mUICC) system.[17, 18] Index date was set as the date of diagnosis.

Treatment modalities

Initial treatments used in the two cohorts consisted of the following: surgical resection, liver transplantation, local ablation therapy (LAT), transarterial chemoembolization/radioembolization (TACE/TARE), radiotherapy, systemic therapy, and best supportive care. These treatment modalities were further categorized as curative treatment (surgical resection, liver transplantation, and LAT), non-curative treatment (TACE/TARE, radiotherapy, and systemic therapy), and best supportive care. TARE and radiotherapy were excluded from the treatment options in the subcohort analysis as they are currently not standardized as primary treatment options in the BCLC recommendations. In principle, the medical, surgical, and interventional procedures for HCC carried out by Korean clinicians were based on the Korean Liver Cancer Association's own practice guidelines internationally recommended for use without modification.[4, 19, 20]

Outcomes

The primary outcome of this study was overall survival. Death certificate data were accessed from the national statistical data collected by the Ministry of Government Administration and Home Affairs in South Korea, and patients who were recorded as alive without specified follow-up date were in all cases labelled with the last evaluation date of a patient diagnosed in the same year. Overall survival outcomes according to sex, liver function, mUICC staging, and type of initial treatment were also obtained.

Although the BCLC staging system is designed to guide the choice of treatment for each stage in accordance with AASLD and EASL practice guidelines, primary treatment of HCC in clinical practice varies widely among patients of the same stage due to differences in underlying liver function and tumor features.[5, 21] Therefore, patients with preserved liver function (Child-Pugh class A) who received the BCLC-recommended treatment options for each stage (BCLC stage 0 or A, single tumor: surgical resection, BCLC stage A with 3 or less nodules each up to 3 cm: LAT, BCLC stage B: TACE, BCLC stage C: systemic therapy),[17] and patients with any degree of liver function who received a liver transplant according to the Milan criteria were further grouped together for the subcohort analysis. The overall survival of these sub-cohorts was then compared to evaluate whether there were differences between the two cohorts even in patients treated according to the same criteria.[2, 3]

Statistical analysis

With regard to baseline characteristics, differences in the distribution of categorical variables were analyzed by the Chi-square test and differences between continuous variables were analyzed by Student's *t*-test or the Wilcoxon rank-sum test. Multivariable Cox proportional hazards models with 95% confidence intervals (CI) were used to assess overall survival, and survival curves were estimated using the Kaplan-Meier method and log-rank test. Because of the retrospective nature of the study, missing data were handled in one or other of two ways: either by analysis with missing data substituted, using the multiple imputation technique, or analysis with missing data classified as a category. Multiple imputation by Markov Chain Monte Carlo methods was used to fill-out incomplete baseline variables, on the assumption that data were missing at random,[22] while interaction analysis was used to evaluate whether the effect of the registry was different within subgroups (sex, liver function, mUICC staging, type of initial treatment).

All statistical analyses were conducted using SAS version 9.4 (SAS Institute Inc, Cary, NC, USA). Two-sided *P*-values ≤0.05 were considered statistically significant.

Results

Study population

Between January 2008 and December 2018, a total of 16,781 patients newly diagnosed with HCC were registered in the KPLCR database (multicenter cohort), and 15,707 patients were recorded in the AMC HCC registry (single-center cohort). After applying the exclusion criteria, a total of 32,098 patients (16,443 patients in the multicenter cohort and 15,655 patients in the single-center cohort) were included in the study (Figure 1).

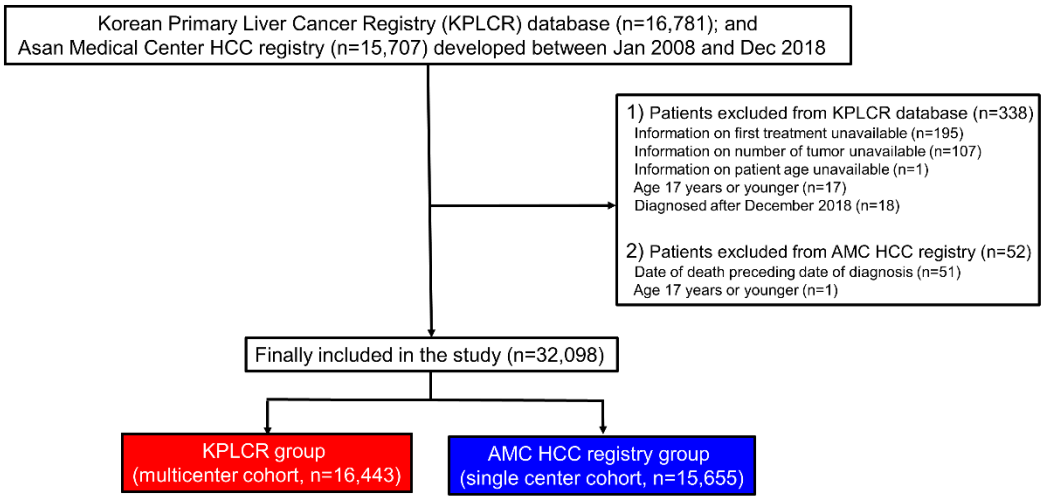


Figure 1. Patient flowchart of the study population

Abbreviations: AMC, Asan Medical Center; HCC, hepatocellular carcinoma; KPLCR, Korean Primary Liver Cancer Registry

Baseline characteristics of the two cohorts are presented in Table 1 and Table 2. Mean ages at diagnosis were 57.7 years (standard deviation [SD], 10.4) and 61.1 years (SD, 11.5) in the single-center and multicenter cohorts, respectively. The single-center cohort had a higher proportion of early-stage patients than the multicenter cohort according to BCLC. Consequently, the use of curative treatment modalities was higher in the single-center cohort, and the use of best supportive care lower.

TABLE 1. Baseline characteristics of the study populations*

Variable	Single-center cohort (n=15,655)	Multicenter cohort (n=16,443)	*P*-value
Age (years)	57.7 ± 10.4	61.1 ± 11.5	<0.001
Male	12,690 (81.1%)	13,045 (79.3%)	<0.001
Body mass index (kg/㎡)	24.3 ± 3.3	24.0 ± 3.4	<0.001
Diabetes mellitus	3,314 (22.0%)	5,779 (40.0%)	<0.001
Hypertension	4,564 (30.2%)	4,347 (33.4%)	<0.001
Hepatitis B†	10,622 (73.1%)	9,879 (62.3%)	<0.001
Hepatitis C‡	1,410 (10.4%)	1,883 (12.7%)	<0.001
mUICC staging			
Stage I	2,626 (16.8%)	2,532 (15.4%)	<0.001
Stage II	6,176 (39.5%)	6,168 (37.6%)	
Stage III	4,712 (30.1%)	4,147 (25.3%)	
Stage IVA	1,291 (8.3%)	1,920 (11.7%)	
Stage IVB	850 (5.4%)	1,627 (9.9%)	
BCLC staging			
Stage 0	2,572 (16.4%)	1,312 (9.3%)	<0.001
Stage A	6,719 (42.9%)	3,655 (25.9%)	
Stage B	2,293 (14.7%)	2,722 (19.3%)	
Stage C	3,563 (22.8%)	5,402 (38.3%)	
Stage D	508 (3.2%)	1,013 (7.2%)	
Child-Pugh class			<0.001
Class A	12,126 (78.0%)	11,476 (73.1 %)	
Class B	2,904 (18.7%)	3,469 (22.1%)	
Class C	510 (3.3%)	747 (4.8%)	
MELD score	8 (7–10)	8 (7–11)	<0.001
Type of initial treatment			
Curative§	6,586 (42.1%)	5,282 (32.1%)	<0.001
Non-curative¶	7,626 (48.7%)	8,070 (49.1%)	
Best supportive care	1,443 (9.2%)	3,091 (18.8%)	
Initial treatment modality			
Surgical resection	5,162 (33.0%)	3,304 (20.1%)	<0.001
Liver transplantation	211 (1.3%)	156 (0.9%)	
LAT	1,213 (7.7%)	1,822 (11.1%)	
TACE/TARE	6,825 (43.6%)	6,839 (41.6%)	
Radiotherapy	186 (1.2%)	245 (1.5%)	

Systemic therapy	615 (3.9%)	986 (6.0%)
Best supportive care	1,443 (9.2%)	3,091 (18.8%)

Data are presented as mean ± standard deviation, median (interquartile range), or frequency (proportion).

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; LAT, local ablation therapy; MELD, Model for End-stage Liver Disease; mUICC, modified Union for International Cancer Control; TACE, transarterial chemoembolization; TARE, transarterial radioembolization.

* Missing data was excluded from the analysis.

† Hepatitis B was defined as any of the following: positive hepatitis B surface antigen, positive viral titer, or previous history of antiviral therapy.

‡ Hepatitis C was defined as any of the following: positive hepatitis C antibody, positive viral titer, or previous history of antiviral therapy.

§ Curative treatment was defined as surgical resection, liver transplantation, and local ablation therapy.

¶ Non-curative treatment was defined as TACE/TARE, radiotherapy, and systemic therapy.

TABLE 2. Baseline characteristics of the study populations (missing data classified as a category)

Variable	Single-center cohort (n=15,655)	Multicenter cohort (n=16,443)	*P*-value
Age (years)	57.7 ± 10.4	61.1 ± 11.5	<0.001
Male	12,690 (81.1%)	13,045 (79.3%)	<0.001
Body mass index (kg/㎡)	24.3 ± 3.3	24.0 ± 3.4	<0.001
Diabetes mellitus	3,314 (21.2%)	5,779 (35.1%)	<0.001
Unknown	615 (3.9%)	2,002 (12.2%)	
Hypertension	4,564 (29.2%)	4,347 (26.4%)	<0.001
Unknown	541 (3.5%)	3,434 (20.9%)	
Hepatitis B†	10,622 (67.8%)	9,879 (60.1%)	<0.001
Unknown	1,125 (7.2%)	592 (3.6%)	
Hepatitis C‡	1,410 (9.0%)	1,883 (11.5%)	<0.001
Unknown	2,126 (13.6%)	1,622 (9.9%)	
mUICC staging			
Stage I	2,626 (16.8%)	2,532 (15.4%)	<0.001
Stage II	6,176 (39.5%)	6,168 (37.5%)	
Stage III	4,712 (30.1%)	4,147 (25.2%)	
Stage IVA	1,291 (8.2%)	1,920 (11.7%)	
Stage IVB	850 (5.4%)	1,627 (9.9%)	
Unknown	0 (0.0%)	49 (0.3%)	
BCLC staging			
Stage 0	2,572 (16.4%)	1,312 (9.0%)	<0.001
Stage A	6,719 (42.9%)	3,655 (22.2%)	
Stage B	2,293 (14.6%)	2,722 (16.6%)	
Stage C	3,563 (22.8%)	5,402 (32.9%)	
Stage D	508 (3.2%)	1,013 (6.2%)	
Unknown	0 (0.0%)	2,339 (14.2%)	
Child-Pugh class			<0.001
Class A	12,126 (77.5%)	11,476 (69.8%)	
Class B	2,904 (18.5%)	3,469 (21.1%)	
Class C	510 (3.3%)	747 (4.5%)	
Unknown	115 (0.7%)	751 (4.6%)	
MELD score	8 (7-10)	8 (7-11)	<0.001
Type of initial treatment			
Curative§	6,586 (42.1%)	5,282 (32.1%)	<0.001

Non-curative¶	7,626 (48.7%)	8,070 (49.1%)	
Best supportive care	1,443 (9.2%)	3,091 (18.8%)	
Initial treatment modality			
Surgical resection	5,162 (33.0%)	3,304 (20.1%)	<0.001
Liver transplantation	211 (1.3%)	156 (0.9%)	
LAT	1,213 (7.7%)	1,822 (11.1%)	
TACE/TARE	6,825 (43.6%)	6,839 (41.6%)	
Radiotherapy	186 (1.2%)	245 (1.5%)	
Systemic therapy	615 (3.9%)	986 (6.0%)	
Best supportive care	1,443 (9.2%)	3,091 (18.8%)	

Data are presented as mean ± standard deviation, median (interquartile range), or frequency (proportion).

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; LAT, local ablation therapy; MELD, Model for End-stage Liver Disease; mUICC, modified Union for International Cancer Control; TACE, transarterial chemoembolization; TARE, transarterial radioembolization.

† Hepatitis B was defined as any of the following: positive hepatitis B surface antigen, positive viral titer, or previous history of antiviral therapy.

‡ Hepatitis C was defined as any of the following: positive hepatitis C antibody, positive viral titer, or previous history of antiviral therapy.

§ Curative treatment was defined as surgical resection, liver transplantation, and LAT.

¶ Non-curative treatment was defined as TACE/TARE, radiotherapy, and systemic therapy.

Distribution of liver function by initial treatment

The distribution of liver function according to Child-Pugh class was identified for each initial treatment modality to assess any differences in distribution between the two cohorts (Table 3). There was no significant difference among the patients who received LAT, whereas the multicenter cohort had a significantly higher proportion of patients with Child-Pugh class A liver function than the single-center cohort among those who received liver transplants (44.4% vs. 30.0%), radiotherapy (58.2% vs 39.7%), systemic therapy (61.8% vs 55.1%), and best supportive care (40.5% vs. 29.6%) (*Ps*<0.001 for all comparisons).

TABLE 3. Distribution of liver function by initial treatment*

Initial treatment modality Child-Pugh class	Single-center cohort (n=15,540)	Multicenter cohort (n=15,692)	*P*-value
Surgical resection	5,126 (33.0%)	3,240 (20.6%)	0.02
Class A	4,902 (95.6%)	3,098 (95.6%)	
Class B	208 (4.1%)	138 (4.3%)	
Class C	16 (0.3%)	4 (0.1%)	
Liver transplantation	210 (1.4%)	153 (1.0%)	0.01
Class A	63 (30.0%)	68 (44.4%)	
Class B	94 (44.8%)	49 (32.0%)	
Class C	53 (25.2%)	36 (23.5%)	
LAT	1,187 (7.6%)	1,766 (11.2%)	0.47
Class A	980 (82.6%)	1,482 (83.9%)	
Class B	186 (15.7%)	249 (14.1%)	
Class C	21 (1.8%)	35 (2.0%)	
TACE/TARE	6,781 (43.6%)	6,527 (41.6%)	<0.001
Class A	5,344 (78.8%)	4,963 (76.0%)	
Class B	1,364 (20.1%)	1,413 (21.7%)	
Class C	73 (1.1%)	151 (2.3%)	
Radiotherapy	184 (1.2%)	237 (1.5%)	<0.001
Class A	73 (39.7%)	138 (58.2%)	
Class B	88 (47.8%)	90 (38.0%)	
Class C	23 (12.5%)	9 (3.8%)	
Systemic therapy	615 (4.0%)	946 (6.0%)	0.002
Class A	339 (55.1%)	585 (61.8%)	
Class B	251 (40.8%)	334 (35.3%)	
Class C	25 (4.1%)	27 (2.9%)	
Best supportive care	1,437 (9.2%)	2,823 (18.0%)	<0.001
Class A	425 (29.6%)	1,142 (40.5%)	
Class B	713 (49.6%)	1,196 (42.4%)	
Class C	299 (20.8%)	485 (17.2%)	

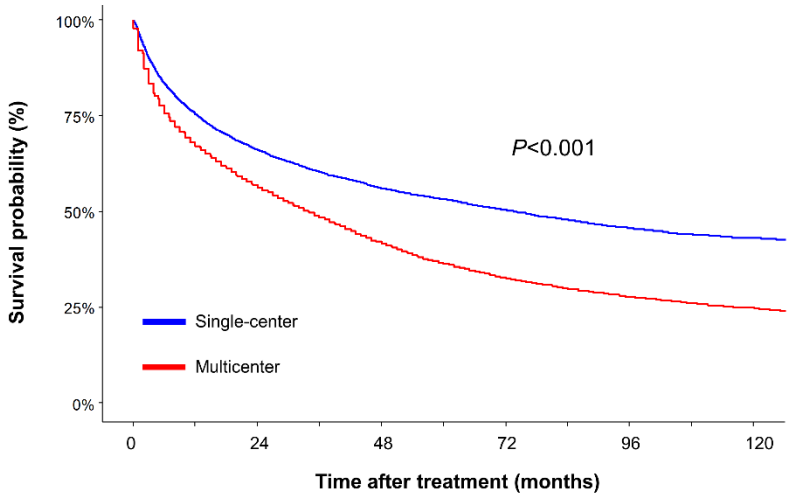
Data are presented as frequency (proportion).

Abbreviations: LAT, local ablation therapy; TACE, transarterial chemoembolization; TARE, transarterial radioembolization.

* Missing data was excluded from the analysis.

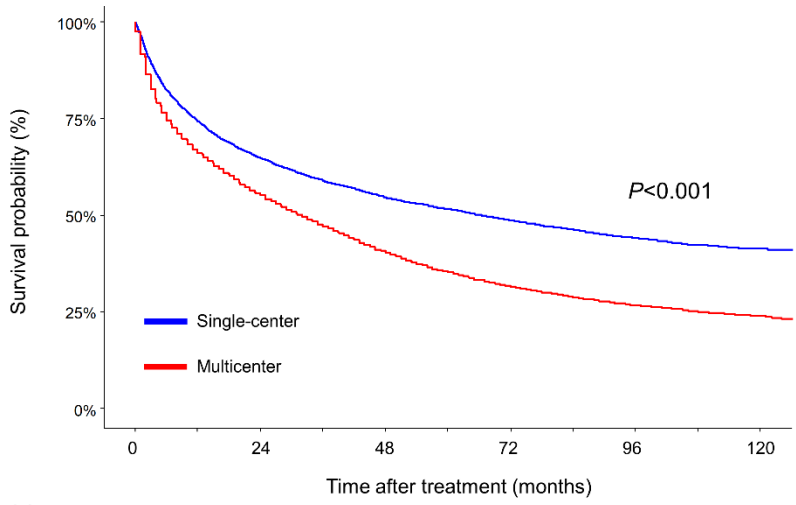
Survival outcomes

The median follow-up duration of single-center and multicenter cohort was 36.2 (interquartile range [IQR]=9.7–66.9) and 30.0 (IQR=6.1–60.0) months, respectively. The single-center cohort had significantly higher overall survival than the multicenter cohort, with median survival times of 73.6 (95% CI=69.6–77.5) and 34.0 (95% CI=33.0–35.0) months, respectively (Figure 2, *P*<0.001 by log-rank test). This finding was consistent regardless of sex, liver function according to Child-Pugh class, and mUICC staging (Figures 3, 4, 5, *Ps*<0.001 for all comparisons).



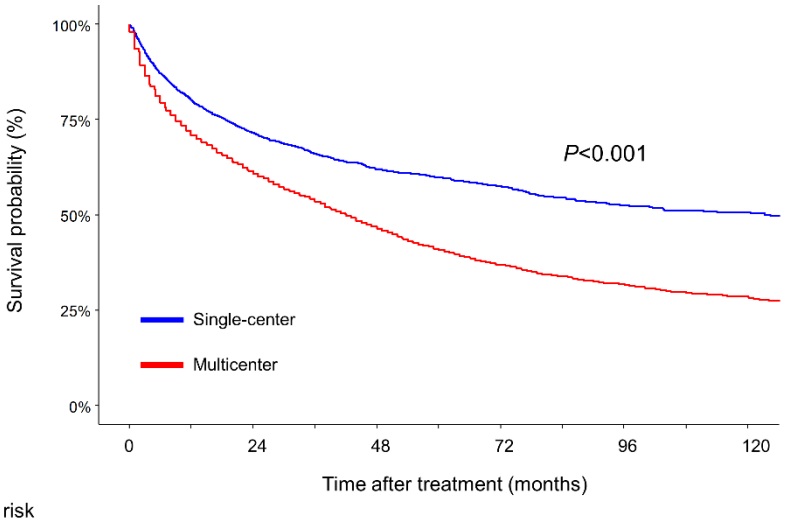
Number at risk						
Single-center	15,655	9,484	6,205	3,566	2,012	749
Multicenter	16,443	9,113	4,745	3,683	2,636	1,500

Figure 2. Kaplan–Meier estimates of overall survival in the two cohorts



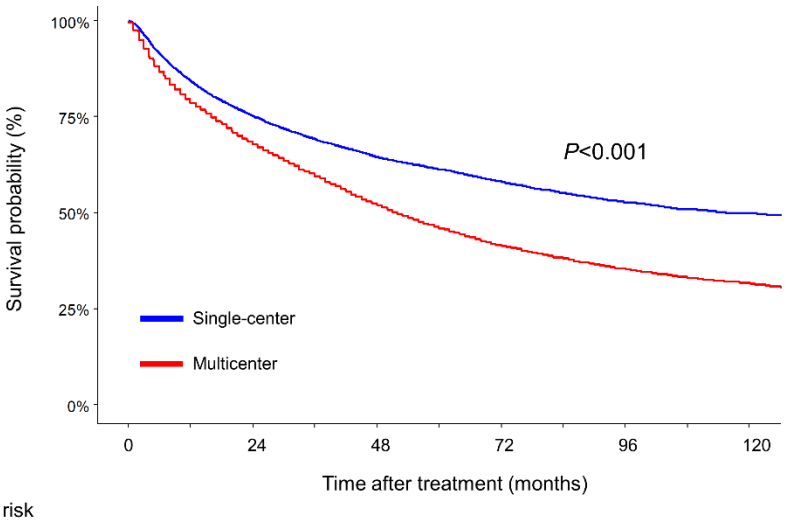
Number at risk						
Single-center	12,690	7,553	4,928	2,810	1,585	571
Multicenter	13,045	7,095	3,649	2,821	2,008	1,141

Figure 3A. Kaplan-Meier estimates of overall survival according to male sex



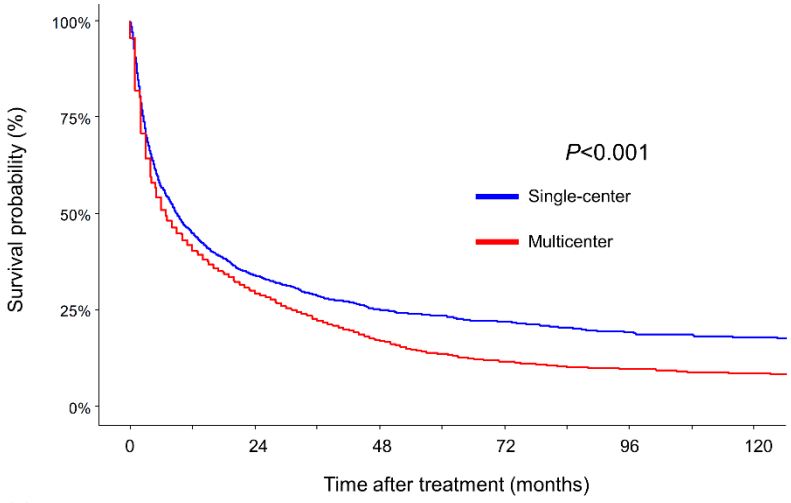
Number at risk						
Single-center	2,965	1,931	1,277	756	427	178
Multicenter	3,398	2,018	1,096	862	628	359

Figure 3B. Kaplan-Meier estimates of overall survival according to female sex



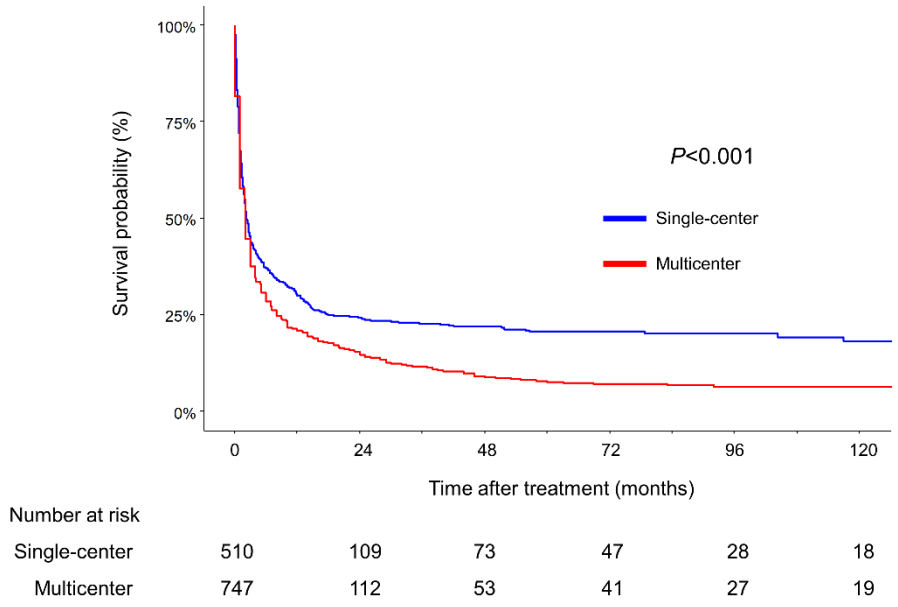
Number at risk						
Single-center	12,126	8,407	5,560	3,186	1,777	636
Multicenter	11,476	7,662	4,043	3,191	2,291	1,275

Figure 4A. Kaplan-Meier estimates of overall survival according to Child-Pugh class A



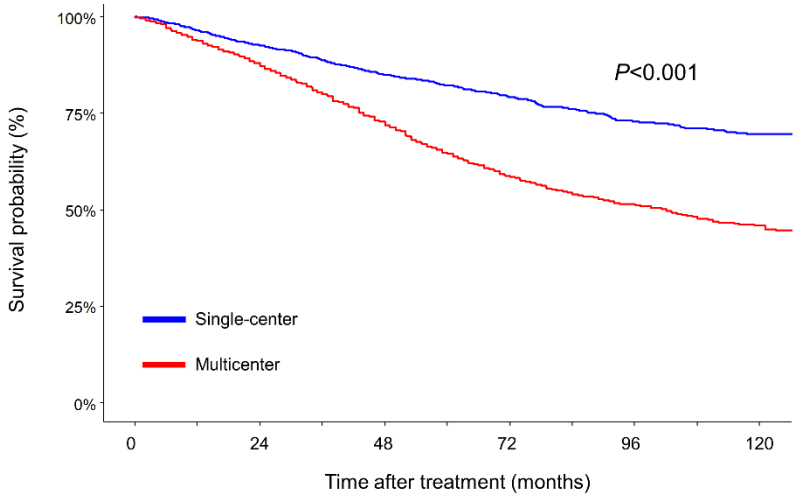
Number at risk						
Single-center	2,904	879	508	297	184	78
Multicenter	3,469	1,018	455	307	214	132

Figure 4B. Kaplan-Meier estimates of overall survival according to Child-Pugh class B



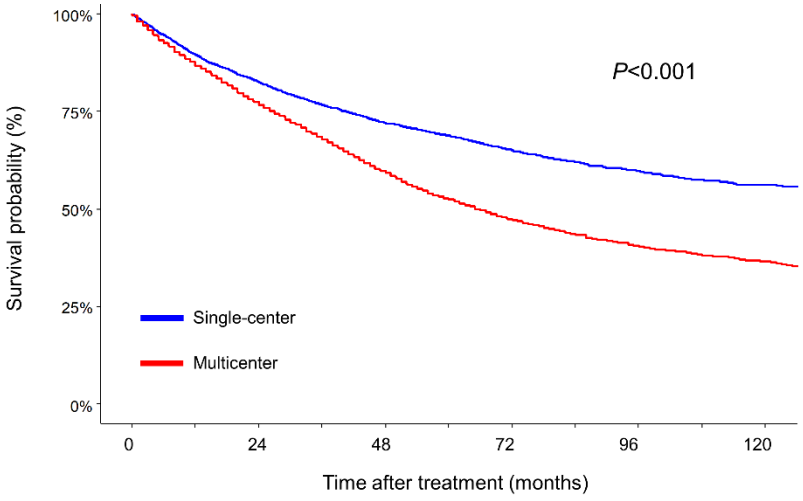
Number at risk						
Single-center	510	109	73	47	28	18
Multicenter	747	112	53	41	27	19

Figure 4C. Kaplan-Meier estimates of overall survival according to Child-Pugh class C



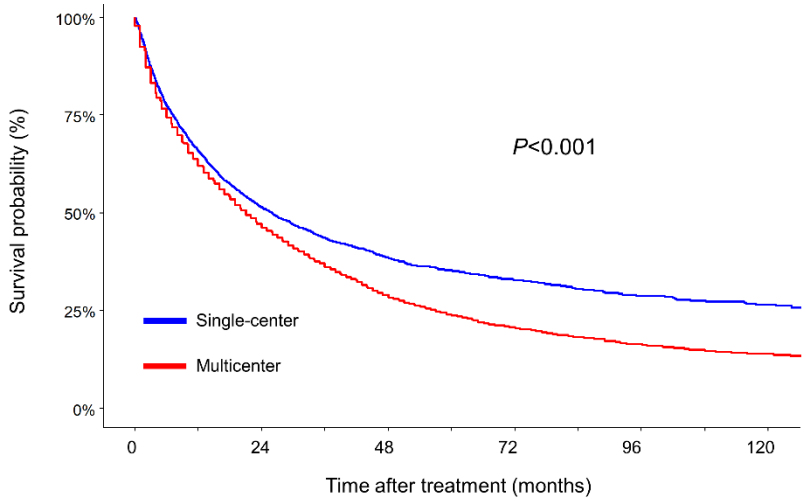
Number at risk						
Single-center	2,626	2,269	1,615	976	596	223
Multicenter	2,532	2,152	1,210	973	719	391

Figure 5A. Kaplan-Meier estimates of overall survival according to modified UICC stage I



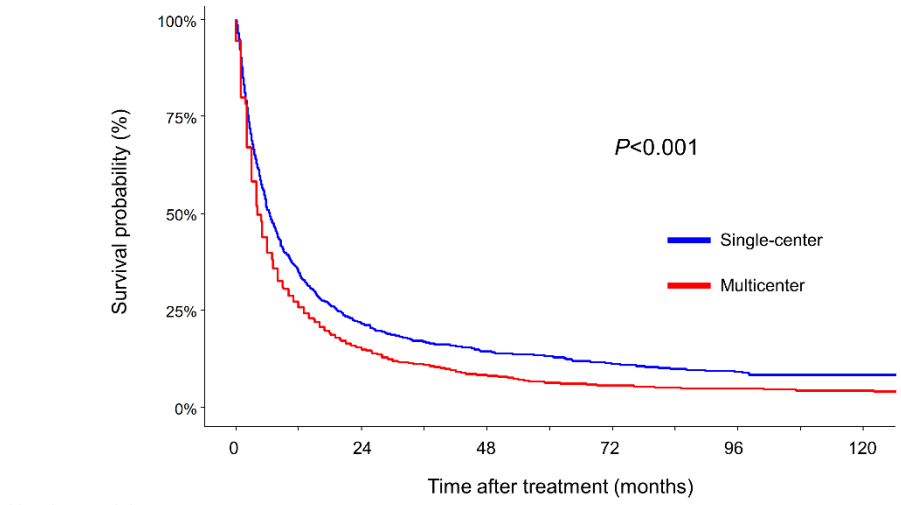
Number at risk						
Single-center	6,176	4,694	3,170	1,797	1,011	385
Multicenter	6,168	4,626	2,515	1,990	1,445	840

Figure 5B. Kaplan-Meier estimates of overall survival according to modified UICC stage II



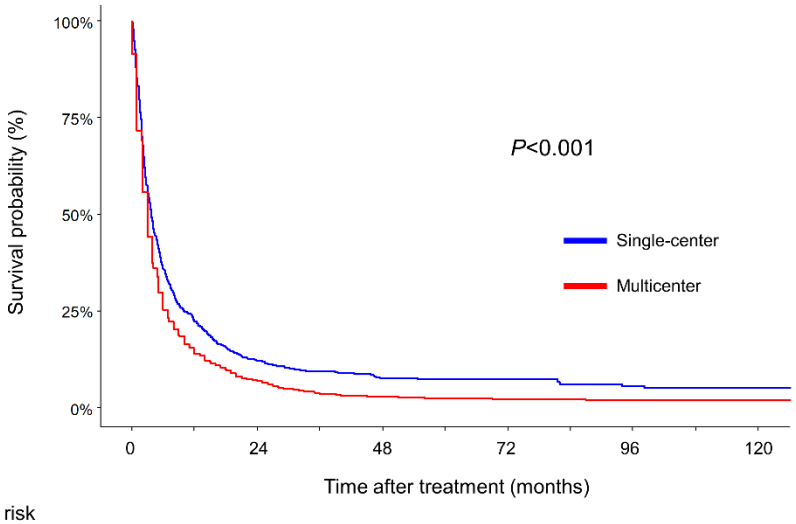
Number at risk						
Single-center	4,712	2,191	1,250	699	360	115
Multicenter	4,147	1,914	850	602	386	214

Figure 5C. Kaplan-Meier estimates of overall survival according to modified UICC stage III



Number at risk						
Single-center	1,291	248	136	78	34	19
Multicenter	1,920	294	123	83	57	32

Figure 5D. Kaplan-Meier estimates of overall survival according to modified UICC stage IVA



Number at risk						
Single-center	850	82	34	16	11	7
Multicenter	1,627	114	40	29	24	19

Figure 5E. Kaplan-Meier estimates of overall survival according to modified UICC stage IVB

In univariate analysis, the multicenter cohort was associated with a significantly higher risk of mortality compared to the single-center cohort (hazards ratio [HR]=1.55, 95% CI=1.50–1.59, P<0.001). Multivariable analysis also showed significantly higher risk of death in the multicenter cohort after adjustment for cancer variables and patient demographics (adjusted hazards ratio [aHR]=1.16, 95% CI=1.13–1.20, P<0.001) (Table 4). Similar results were shown with missing data classified as a category, presented in Table 5.

TABLE 4. Cox regression analysis of factors associated with mortality in the entire cohorts*

Variable	Univariate analysis		Multivariable analysis with multiple imputation	
	HR (95% CI)	*P*-value	Adjusted HR (95% CI)	*P*-value
Cohort				
Single-center	1 (reference)		1 (reference)	
Multicenter	1.55 (1.50–1.59)	<0.001	1.16 (1.13–1.20)	<0.001
Age ≥ 60 years	1.22 (1.19–1.26)	<0.001	1.13 (1.10–1.17)	<0.001
Female (vs. Male)	0.83 (0.80–0.87)	<0.001	0.93 (0.89–0.96)	<0.001
Hepatitis B†	0.77 (0.75–0.80)	<0.001	0.95 (0.91–0.98)	0.003
Hepatitis C‡	1.18 (1.13–1.24)	<0.001	1.06 (1.00–1.11)	0.04
mUICC staging				
Stage I	1 (reference)	<0.001	1 (reference)	<0.001
Stage II	1.56 (1.47–1.65)	<0.001	1.50 (1.42–1.58)	<0.001
Stage III	3.81 (3.60–4.03)	<0.001	2.78 (2.62–2.94)	<0.001
Stage IVA	9.32 (8.76–9.92)	<0.001	5.55 (5.20–5.92)	<0.001
Stage IVB	14.59 (13.67–15.58)	<0.001	8.07 (7.54–8.63)	<0.001
Child-Pugh class				
Class A	1 (reference)	<0.001	1 (reference)	<0.001
Class B	3.09 (2.99–3.20)	<0.001	1.94 (1.87–2.01)	<0.001
Class C	5.05 (4.74–5.37)	<0.001	3.12 (2.92–3.34)	<0.001
Type of initial treatment				
Curative§	1 (reference)	<0.001	1 (reference)	<0.001
Non-curative¶	3.64 (3.50–3.78)	<0.001	2.39 (2.30–2.49)	<0.001
Best supportive care	12.89 (12.31–13.50)	<0.001	5.71 (5.42–6.01)	<0.001
Initial treatment modality				
Surgical resection	1 (reference)	<0.001		
Liver transplantation	0.76 (0.61–0.95)	0.02		
LAT	1.43 (1.33–1.54)	<0.001		
TACE/TARE	3.55 (3.39–3.72)	<0.001		
Radiotherapy	8.75 (7.83–9.78)	<0.001		
Systemic therapy	14.82 (13.85–15.85)	<0.001		
Best supportive care	14.70 (13.95–15.50)	<0.001		

Abbreviations: CI, confidence interval; HR, hazard ratio; LAT, local ablation therapy; mUICC, modified Union for International Cancer Control; TACE, transarterial chemoembolization; TARE, transarterial radioembolization.

* Missing data was imputed.

† Hepatitis B was defined as any of the following: positive hepatitis B surface antigen, positive viral titer, or previous history of antiviral therapy.

‡ Hepatitis C was defined as any of the following: positive hepatitis C antibody, positive viral titer, or previous history of antiviral therapy.

§ Curative treatment was defined as surgical resection, liver transplantation, and local ablation therapy.

¶ Non-curative treatment was defined as TACE/TARE, radiotherapy, and systemic therapy.

TABLE 5. Cox regression analysis of factors associated with mortality in the entire cohort (missing data classified as a category)

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P*-value	Adjusted HR (95% CI)	*P*-value
Cohort				
Single-center	1 (reference)		1 (reference)	
Multicenter	1.55 (1.50–1.59)	<0.001	1.17 (1.13–1.20)	<0.001
Age ≥ 60 years	1.22 (1.19–1.26)	<0.001	1.14 (1.10–1.18)	<0.001
Female (vs. Male)	0.83 (0.80–0.87)	<0.001	0.93 (0.89–0.96)	<0.001
Hepatitis B†	0.77 (0.75–0.79)	<0.001	0.96 (0.92–0.99)	0.02
Unknown	0.70 (0.65–0.75)	<0.001	0.92 (0.85–0.99)	0.03
Hepatitis C‡	1.18 (1.13–1.24)	<0.001	1.07 (1.02–1.13)	0.006
Unknown	0.83 (0.79–0.87)	<0.001	0.98 (0.92–1.03)	0.37
Modified UICC staging				
Stage I	1 (reference)	<0.001	1 (reference)	<0.001
Stage II	1.56 (1.47–1.65)	<0.001	1.49 (1.41–1.58)	<0.001
Stage III	3.81 (3.60–4.02)	<0.001	2.75 (2.60–2.91)	<0.001
Stage IVA	9.31 (8.75–9.91)	<0.001	5.49 (5.15–5.85)	<0.001
Stage IVB	14.57 (13.65–15.56)	<0.001	7.96(7.44–8.52)	<0.001
Unknown	6.87 (5.09–9.27)	<0.001	3.61 (2.67–4.88)	<0.001
Child-Pugh class				
Class A	1 (reference)	<0.001	1 (reference)	<0.001
Class B	3.09 (2.99–3.19)	<0.001	2.01 (1.94–2.08)	<0.001
Class C	5.04 (4.73–5.37)	<0.001	3.18 (2.98–3.39)	<0.001
Unknown	1.86 (1.72–2.02)	<0.001	1.19 (1.10–1.30)	<0.001
Type of initial treatment				
Curative§	1 (reference)	<0.001	1 (reference)	<0.001
Non-curative¶	3.64 (3.50–3.78)	<0.001	2.39 (2.30–2.49)	<0.001
Best supportive care	12.89 (12.31–13.50)	<0.001	5.71 (5.42–6.01)	<0.001
Initial treatment modality				
Surgical resection	1 (reference)	<0.001		
Liver transplantation	0.76 (0.61–0.95)	0.02		
LAT	1.43 (1.33–1.54)	<0.001		
TACE/TARE	3.55 (3.39–3.72)	<0.001		
Radiotherapy	8.75 (7.83–9.78)	<0.001		
Systemic therapy	14.82 (13.85–15.85)	<0.001		
Best supportive care	14.70 (13.95–15.50)	<0.001		

Abbreviations: CI, Confidence interval; HR, Hazard ratio; LAT, Local ablation therapy; mUICC, Union for International Cancer Control; TACE, Transarterial chemoembolization; TARE, Transarterial radioembolization.

† Hepatitis B was defined as any of the following: positive hepatitis B surface antigen, positive viral titer, or previous history of antiviral therapy.

‡ Hepatitis C was defined as any of the following: positive hepatitis C antibody, positive viral titer, or previous history of antiviral therapy.

§ Curative treatment was defined as surgical resection, liver transplantation, and LAT.

¶ Non-curative treatment was defined as TACE/TARE, radiotherapy, and systemic therapy.

Comparisons of overall survival in the entire cohorts according to initial treatment yielded variable results (Table 6 and Figures 6 and 7). Multivariable analysis with multiple imputation revealed a higher risk of mortality in the multicenter cohort in patients who received surgical resection (aHR=1.32, 95% CI=1.22–1.44, *P*<0.001), LAT (aHR=1.50, 95% CI=1.32–1.71, *P*<0.001), TACE/TARE (aHR=1.24, 95% CI=1.19–1.29, *P*<0.001), and liver transplantation (aHR=2.10, 95% CI=1.30–3.38, *P*=0.002). Overall, there was a higher risk of death among patients in the multicenter cohort who received curative treatment (aHR=1.48, 95% CI=1.39–1.59, *P*<0.001) or non-curative treatment (aHR=1.22, 95% CI=1.17–1.27, *P*<0.001), and death was significantly lower in patients who received systemic therapy (aHR=0.83, 95% CI=0.74–0.92, *P*=0.001) and best supportive care (aHR=0.85, 95% CI=0.79–0.91, *P*<0.001). Overall survival following radiotherapy as an initial option, however, did not differ significantly between the two cohorts (aHR=1.14, 95% CI=0.91–1.42, *P*=0.25). The results of multivariable analysis with missing data classified as a category are presented in Table 7, and gave similar outcomes.

TABLE 6. Cox regression analysis of risk of mortality by initial treatment in the entire cohorts*

Initial treatment modality	Univariate analysis		Multivariable analysis with multiple imputation†	
	HR (95% CI)	*P*-value	Adjusted HR (95% CI)	*P*-value
Surgical resection				
Single-center (n=5,162)	1 (reference)		1 (reference)	
Multicenter (n=3,304)	1.38 (1.27–1.50)	<0.001	1.32 (1.22 –1.44)	<0.001
Liver transplantation				
Single-center (n=211)	1 (reference)		1 (reference)	
Multicenter (n=156)	2.22 (1.41–3.51)	<0.001	2.10 (1.30–3.38)	0.002
LAT				
Single-center (n=1,213)	1 (reference)		1 (reference)	
Multicenter (n=1,822)	1.64 (1.45–1.87)	<0.001	1.50 (1.32–1.71)	<0.001
TACE/TARE				
Single-center (n=6,825)	1 (reference)		1 (reference)	
Multicenter (n=6,839)	1.25 (1.20–1.31)	<0.001	1.24 (1.19–1.29)	<0.001
Radiotherapy				
Single-center (n=186)	1 (reference)		1 (reference)	
Multicenter (n=245)	1.16 (0.94–1.43)	0.16	1.14 (0.91–1.42)	0.25
Systemic therapy				
Single-center (n=615)	1 (reference)		1 (reference)	
Multicenter (n=986)	0.86 (0.78–0.96)	0.007	0.83 (0.74–0.92)	0.001
Curative treatment‡				
Single-center (n=6,586)	1 (reference)	<0.001	1 (reference)	<0.001
Multicenter (n=5,282)	1.54 (1.44–1.65)	<0.001	1.48 (1.39–1.59)	<0.001
Non-curative treatment§				
Single-center (n=7,626)	1 (reference)	<0.001	1 (reference)	<0.001
Multicenter (n=8,070)	1.28 (1.23–1.33)	<0.001	1.22 (1.17–1.27)	<0.001
Best supportive care				
Single-center (n=1,443)	1 (reference)	<0.001	1 (reference)	<0.001
Multicenter (n=3,091)	0.85 (0.80–0.91)	<0.001	0.85 (0.79–0.91)	<0.001

Abbreviations: CI, confidence interval; HR, hazard ratio; LAT, local ablation therapy ablation; TACE, transarterial chemoembolization; TARE, transarterial radioembolization.

* Missing data was imputed.

† Adjusted for sex, age, hepatitis B, hepatitis C, Child-Pugh class, and modified Union for International Cancer Control (mUICC) staging.

‡ Curative treatment was defined as surgical resection, liver transplantation, and local ablation therapy.

§ Non-curative treatment was defined as TACE/TARE, radiotherapy, and systemic therapy.

TABLE 7. Cox regression analysis of risk of mortality by initial treatment in the entire cohorts (missing data classified as a category)

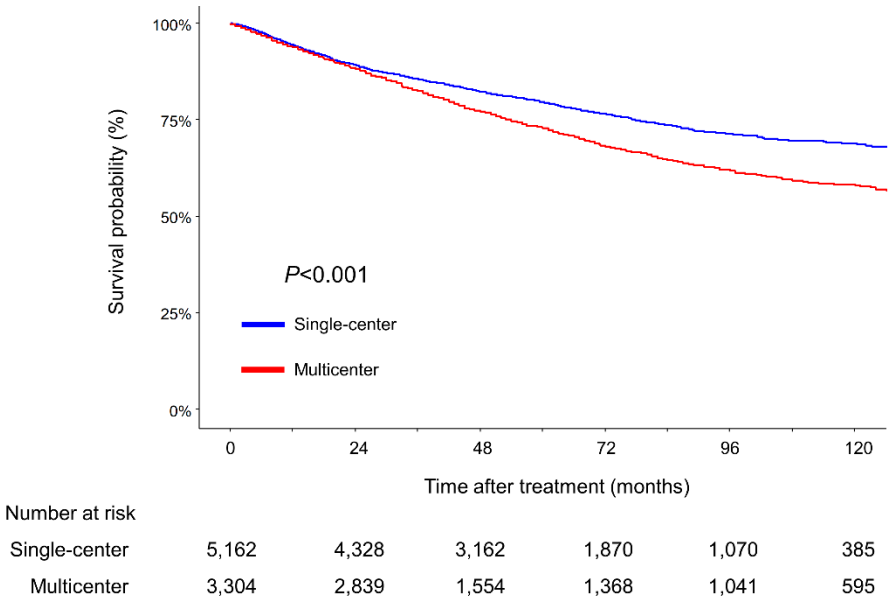
Initial treatment modality	Univariate analysis		Multivariate analysis†	
	HR (95% CI)	*P*-value	Adjusted HR (95% CI)	*P*-value
Surgical resection				
Single-center (n=5,162)	1 (reference)		1 (reference)	
Multicenter (n=3,304)	1.38 (1.27–1.50)	<0.001	1.32 (1.21–1.44)	<0.001
Liver transplantation				
Single-center (n=211)	1 (reference)		1 (reference)	
Multicenter (n=156)	2.22 (1.41–3.51)	<0.001	1.91 (1.18–3.10)	0.009
LAT				
Single-center (n=1,213)	1 (reference)		1 (reference)	
Multicenter (n=1,822)	1.64 (1.45–1.87)	<0.001	1.56 (1.35–1.79)	<0.001
TACE/TARE				
Single-center (n=6,825)	1 (reference)		1 (reference)	
Multicenter (n=6,839)	1.25 (1.20–1.31)	<0.001	1.24 (1.19–1.30)	<0.001
Radiotherapy				
Single-center (n=186)	1 (reference)		1 (reference)	
Multicenter (n=245)	1.16 (0.94–1.43)	0.16	1.14 (0.92–1.43)	0.23
Systemic therapy				
Single-center (n=615)	1 (reference)		1 (reference)	
Multicenter (n=986)	0.86 (0.78–0.96)	0.007	0.82 (0.74–0.92)	0.001
Curative treatment‡				
Single-center (n=6,586)	1 (reference)	<0.001	1 (reference)	<0.001
Multicenter (n=5,282)	1.54 (1.44–1.65)	<0.001	1.50 (1.40–1.60)	<0.001
Non-curative treatment§				
Single-center (n=7,626)	1 (reference)	<0.001	1 (reference)	<0.001
Multicenter (n=8,070)	1.28 (1.23–1.33)	<0.001	1.22 (1.18–1.27)	<0.001
Best supportive care				
Single-center (n=1,443)	1 (reference)	<0.001	1 (reference)	<0.001
Multicenter (n=3,091)	0.85 (0.80–0.91)	<0.001	0.86 (0.80–0.92)	<0.001

Abbreviations: CI, confidence interval; HR, hazard ratio; LAT, local ablation therapy ablation; TACE, transarterial chemoembolization; TARE, transarterial radioembolization.

† Adjusted for sex, age, hepatitis B, hepatitis C, Child-Pugh class, and modified Union for International Cancer Control (mUICC) staging.

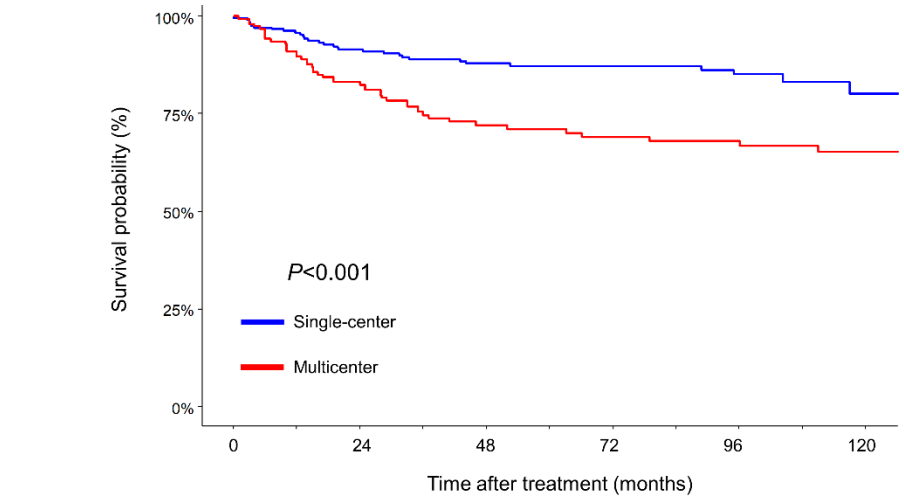
‡ Curative treatment was defined as surgical resection, liver transplantation, and LAT.

§ Non-curative treatment was defined as TACE/TARE, radiotherapy, and systemic therapy.



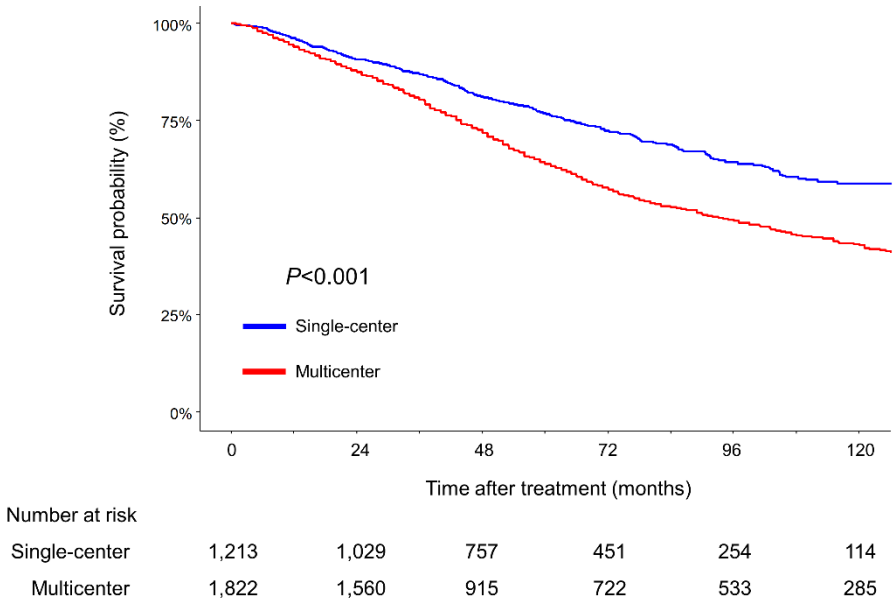
Number at risk						
Single-center	5,162	4,328	3,162	1,870	1,070	385
Multicenter	3,304	2,839	1,554	1,368	1,041	595

Figure 6A. Kaplan-Meier estimates of overall survival following surgical resection



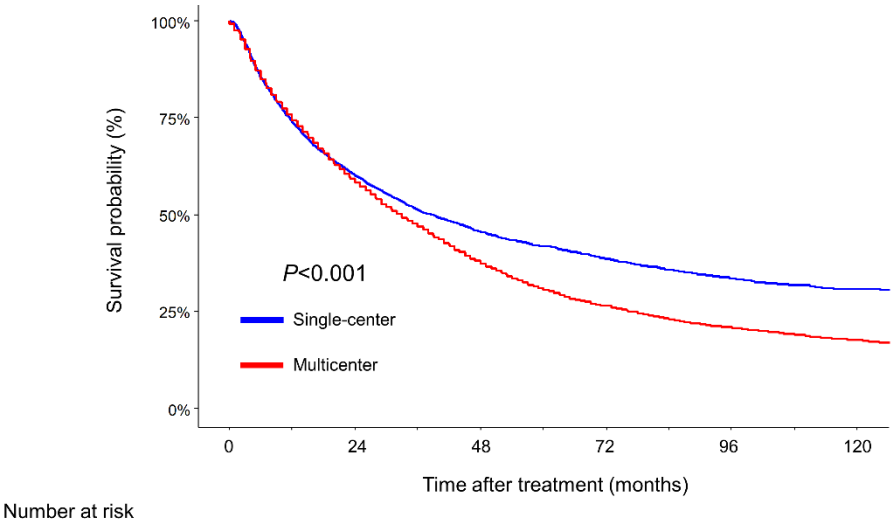
Number at risk						
Single-center	211	188	153	113	72	26
Multicenter	156	125	74	70	53	39

Figure 6B. Kaplan-Meier estimates of overall survival following liver transplantation



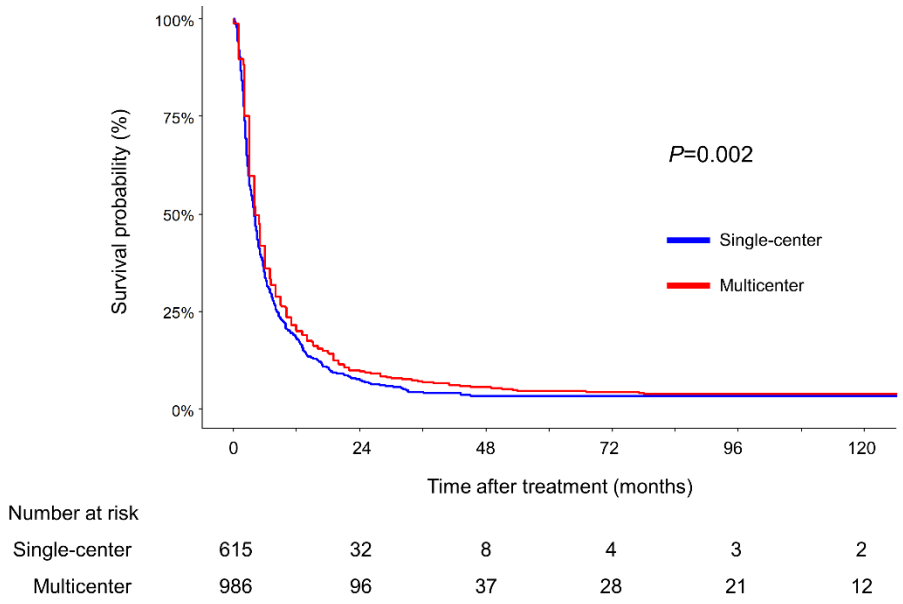
Number at risk						
Single-center	1,213	1,029	757	451	254	114
Multicenter	1,822	1,560	915	722	533	285

Figure 6C. Kaplan-Meier estimates of overall survival following local ablation therapy



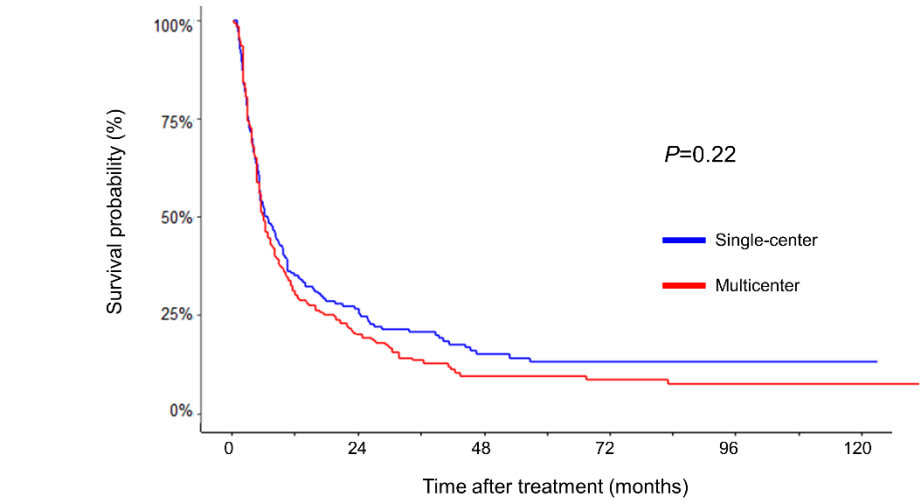
Number at risk						
Single-center	6,825	3,741	2,058	1,088	597	213
Multicenter	6,839	3,974	1,912	1,330	880	513

Figure 6D. Kaplan-Meier estimates of overall survival following transarterial chemoembolization/transarterial radioembolization



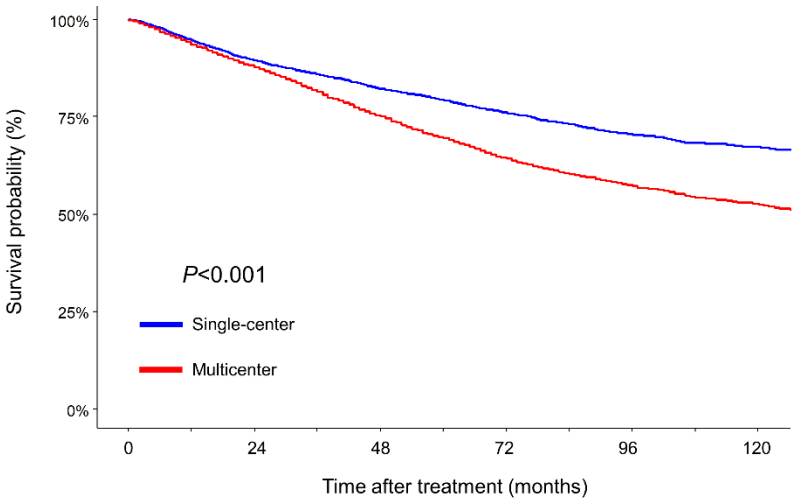
Number at risk						
Single-center	615	32	8	4	3	2
Multicenter	986	96	37	28	21	12

Figure 6E. Kaplan-Meier estimates of overall survival following systemic therapy



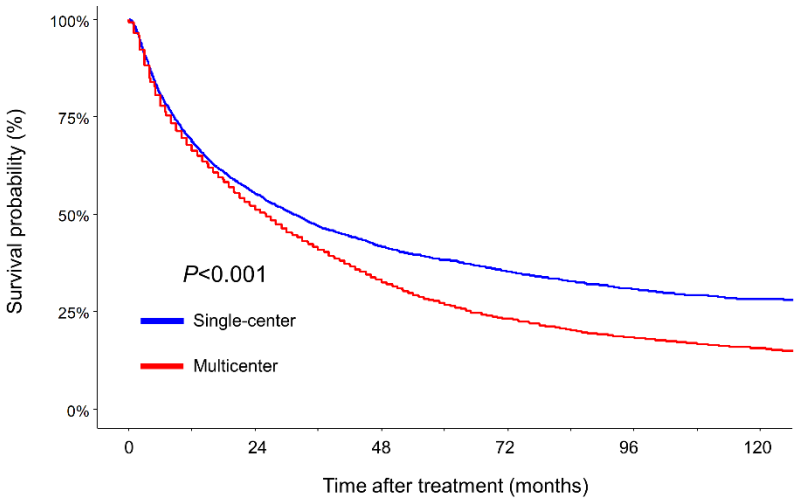
Number at risk						
Single-center	186	44	21	14	6	3
Multicenter	245	55	14	11	8	5

Figure 6F. Kaplan-Meier estimates of overall survival following radiotherapy



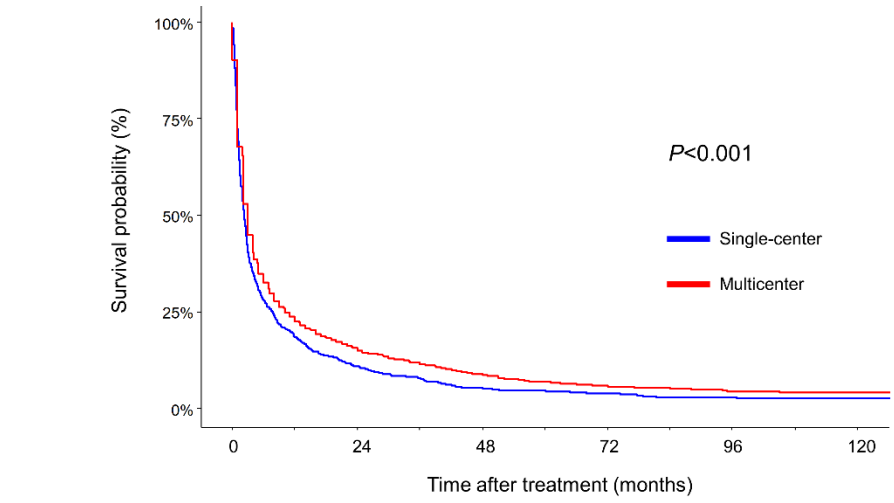
Number at risk						
Single-center	6,586	5,545	4,072	2,434	1,396	525
Multicenter	5,282	4,524	2,543	2,160	1,627	919

Figure 7A. Kaplan-Meier estimates of overall survival following curative treatment



Number at risk						
Single-center	7,626	3,817	2,087	1,106	606	218
Multicenter	8,070	4,125	1,963	1,369	909	530

Figure 7B. Kaplan-Meier estimates of overall survival following non-curative treatment



Number at risk						
Single-center	1,443	122	46	26	10	6
Multicenter	3,091	464	239	154	100	51

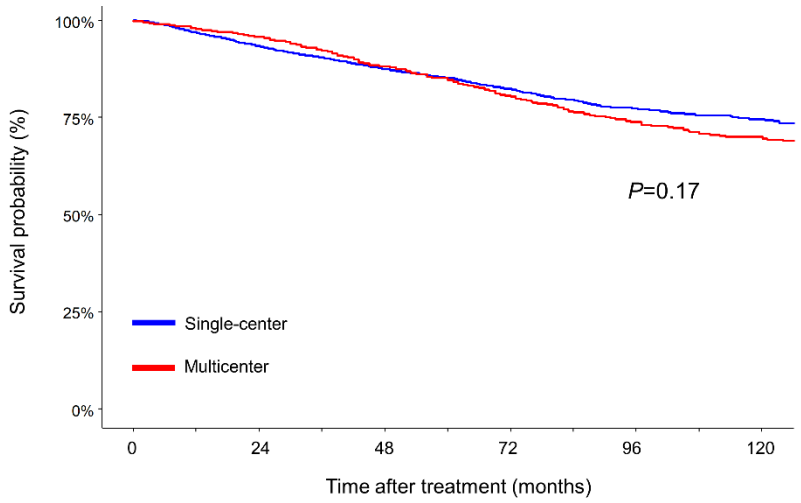
Figure 7C. Kaplan-Meier estimates of overall survival following best supportive care

Subcohort analysis of patients treated according to BCLC guidelines

Subcohort analysis was conducted to further compare survival outcomes between two sub-cohorts (n=2,797 and n=5,151 for multicenter and single-center subsets, respectively) comprised of patients with preserved liver function (Child-Pugh class A) who received treatment according to the BCLC strategy, and patients with any level of liver function who received liver transplants according to the Milan criteria.

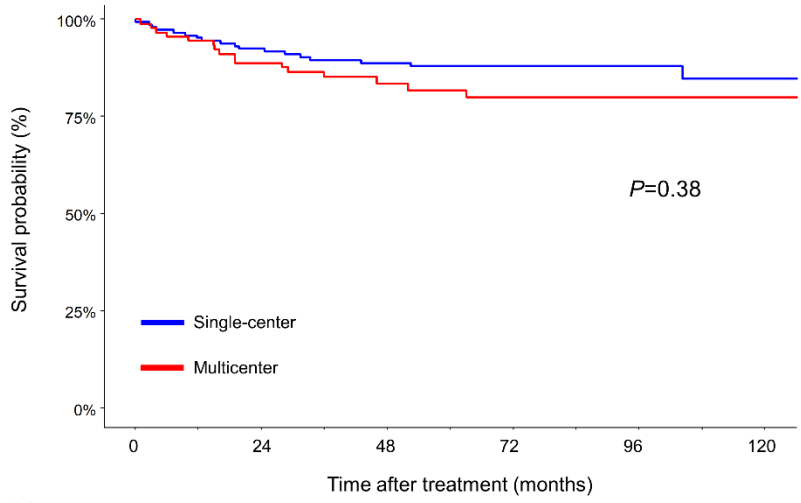
Overall survival did not differ between the two sub-cohorts in patients who received surgical resection (*P*=0.17 by log-rank test; Figure 8A), liver transplants (*P*=0.38; Figure 8B), and systemic therapy (median survival time, 5.5 [IQR=5.0–6.1] and 5.1 [IQR=4.5–6.0] months, respectively, *P*=0.23; Figure 8E). These findings were confirmed in multivariable analysis: risk of mortality among patients with preserved liver function who received surgical resection (aHR=1.07, 95% CI=0.93–1.23, *P*=0.33) or systemic therapy (aHR=0.94, 95% CI=0.81–1.10, *P*=0.44) did not differ between the two cohorts, and for patients who received liver transplants within the Milan criteria (aHR=1.30, 95% CI=0.65–2.60, *P*=0.45) (Table 8).

Among patients with preserved liver function who received either TACE or LAT in accordance with the BCLC treatment strategy, the multicenter subcohort was associated with a higher risk of death than the single-center subcohort in both univariate (HR=1.74, 95% CI=1.50–2.02, *P*<0.001; and HR=1.42, 95% CI=1.07–1.90, *P*=0.02, respectively) and multivariable analyses (aHR=1.72, 95% CI=1.48–2.00, *P*<0.001; and aHR=1.44, 95% CI=1.08–1.92, *P*=0.01, respectively). Similar outcomes were obtained in multivariable analysis with missing data classified as individual categories (Table 9).



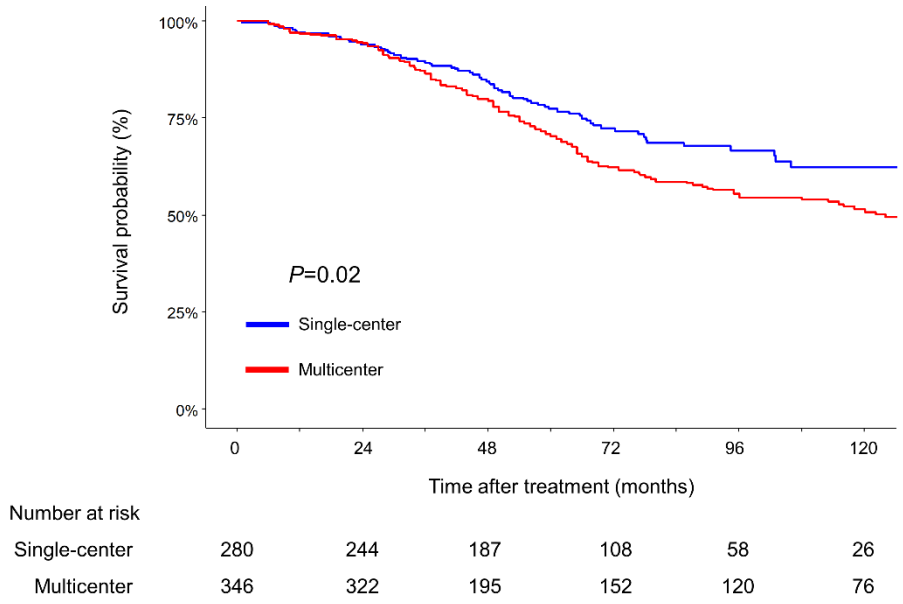
Number at risk						
Single-center	3,771	3,295	2,414	1,418	819	280
Multicenter	1,481	1,393	788	715	548	304

Figure 8A. Kaplan–Meier estimates of overall survival of patients who received surgical resection according to the treatment indications



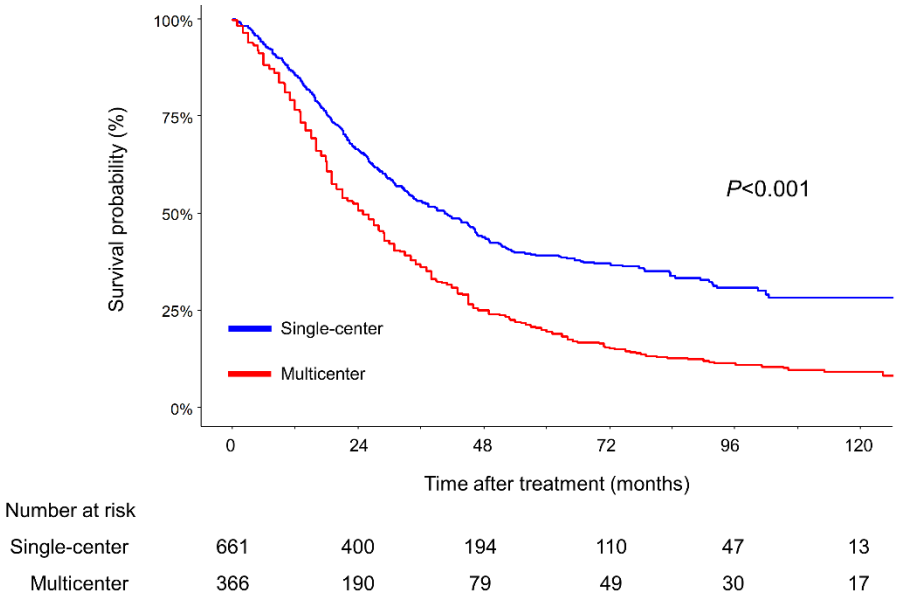
Number at risk						
Single-center	146	132	105	75	48	13
Multicenter	90	77	49	46	35	25

Figure 8B. Kaplan–Meier estimates of overall survival of patients who received liver transplant according to the treatment indications



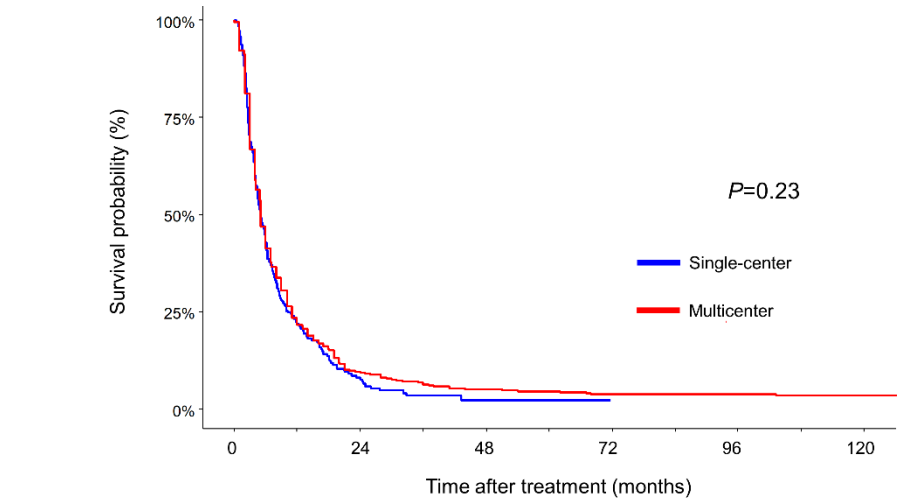
Number at risk						
Single-center	280	244	187	108	58	26
Multicenter	346	322	195	152	120	76

Figure 8C. Kaplan–Meier estimates of overall survival of patients who received local ablation therapy according to the treatment indications



Number at risk						
Single-center	661	400	194	110	47	13
Multicenter	366	190	79	49	30	17

Figure 8D. Kaplan–Meier estimates of overall survival of patients who received transarterial chemoembolization according to the treatment indications



Number at risk						
Single-center	293	15	2	0	0	0
Multicenter	514	49	20	15	13	8

Figure 8E. Kaplan–Meier estimates of overall survival of patients who received systemic therapy according to the treatment indications

TABLE 8. Cox regression analysis of risk of mortality by initial treatment in BCLC-guided sub-cohorts*

Initial treatment modality	Univariate analysis		Multivariable analysis with multiple imputation†	
	HR (95% CI)	*P*-value	Adjusted HR (95% CI)	*P*-value
Surgical resection				
Single-center (n=3,771)	1 (reference)		1 (reference)	
Multicenter (n=1,481)	1.10 (0.96–1.26)	0.17	1.07 (0.93–1.23)	0.33
Liver transplantation				
Single-center (n=146)	1 (reference)		1 (reference)	
Multicenter (n=90)	1.35 (0.69–2.66)	0.38	1.30 (0.65–2.60)	0.45
LAT				
Single-center (n=280)	1 (reference)		1 (reference)	
Multicenter (n=346)	1.42 (1.07–1.90)	0.02	1.44 (1.08–1.92)	0.01
TACE				
Single-center (n=661)	1 (reference)		1 (reference)	
Multicenter (n=366)	1.74 (1.50–2.02)	<0.001	1.72 (1.48–2.00)	<0.001
Systemic therapy				
Single-center (n=293)	1 (reference)		1 (reference)	
Multicenter (n=514)	0.93 (0.80–1.08)	0.33	0.94 (0.81–1.10)	0.44

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; CI, confidence interval; HR, hazard ratio; LAT, local ablation therapy; TACE, transarterial chemoembolization.

* Missing data was imputed.

† Adjusted for sex, age, hepatitis B, hepatitis C, Child-Pugh class, and modified Union for International Cancer Control (mUICC) staging.

TABLE 9. Cox regression analysis of risk of mortality by initial treatment in the BCLC-guided subcohorts (missing data classified as a category)

Initial treatment modality	Univariate analysis		Multivariate analysis†	
	HR (95% CI)	*P*-value	Adjusted HR (95% CI)	*P*-value
Surgical resection				
Single-center (n=3,771)	1 (reference)		1 (reference)	
Multicenter (n=1,481)	1.10 (0.96–1.26)	0.17	1.06 (0.93–1.22)	0.38
Liver transplantation				
Single-center (n=146)	1 (reference)		1 (reference)	
Multicenter (n=90)	1.35 (0.69–2.66)	0.38	1.09 (0.54–2.18)	0.82
LAT				
Single-center (n=280)	1 (reference)		1 (reference)	
Multicenter (n=346)	1.42 (1.07–1.90)	0.02	1.45 (1.08–1.95)	0.01
TACE				
Single-center (n=661)	1 (reference)		1 (reference)	
Multicenter (n=366)	1.74 (1.50–2.02)	<0.001	1.73 (1.49–2.02)	<0.001
Systemic therapy				
Single-center (n=293)	1 (reference)		1 (reference)	
Multicenter (n=514)	0.93 (0.80–1.08)	0.33	0.94 (0.81–1.10)	0.46

Abbreviations: BCLC, Barcelona Clinic Liver Cancer; CI, confidence interval; HR, hazard ratio; LAT, local ablation therapy; TACE, transarterial chemoembolization.

† Adjusted for sex, age, hepatitis B, hepatitis C, Child-Pugh class, and modified Union for International Cancer Control (mUICC) staging.

Subgroup analysis

Interaction analysis performed to evaluate the effect of type of registry in the different subgroups showed that the multicenter cohort was associated with a significantly higher risk of mortality in both sexes, for all degrees of liver function, as well as for all stages of the mUICC system (*Ps*<0.001 for all subgroups) (Table 10). As regards initial treatment modality, subgroups of the multicenter cohort who received curative or non-curative treatment had higher risks of mortality, but overall survival was higher in the subgroup that received best supportive care (*Ps*<0.001 for all).

TABLE 10. Subgroup analysis*

Subgroup	Single-center cohort		Multicenter cohort				
	Cases	Events (%)	Cases	Events (%)	Crude HR (95% CI)†	*P*-value	*P* for interaction
Sex							0.013
Male	12,690	6,126 (48.3%)	13,045	8,893 (68.2%)	1.52 (1.48–1.58)	<0.001	
Female	2,965	1,190 (40.1%)	3,398	2,142 (63.0%)	1.68 (1.57–1.81)	<0.001	
mUICC staging							<0.001
Stage I	2,626	527 (20.1%)	2,532	1,058 (41.8%)	2.13 (1.91–2.36)	<0.001	
Stage II	6,176	1,991 (32.2%)	6,168	3,273 (53.1%)	1.65 (1.56–1.74)	<0.001	
Stage III	4,712	2,954 (62.7%)	4,147	3,267 (78.8%)	1.29 (1.23–1.36)	<0.001	
Stage IVA	1,291	1,100 (85.2%)	1,920	1,807 (94.1%)	1.32 (1.23–1.42)	<0.001	
Stage IVB	850	744 (87.5%)	1,627	1,586 (97.5%)	1.34 (1.23–1.46)	<0.001	
Child-Pugh class							<0.001
Class A	12,126	4,738 (39.1%)	11,476	6,717 (58.5%)	1.55 (1.49–1.61)	<0.001	
Class B	2,904	2,156 (74.2%)	3,469	3,039 (87.6%)	1.26 (1.19–1.33)	<0.001	
Class C	510	391 (76.7%)	747	686 (91.8%)	1.53 (1.35–1.73)	<0.001	
Type of initial treatment							<0.001
Curative‡	6,586	1,523 (23.1%)	5,282	1,945 (36.8%)	1.58 (1.47–1.67)	<0.001	
Non-curative§	7,626	4,533 (59.4%)	8,070	6,222 (77.1%)	1.27 (1.22–1.32)	<0.001	
Best supportive care	1,443	1,260 (87.3%)	3,091	2,868 (92.8%)	0.75 (0.70–0.80)	<0.001	

Abbreviations: CI, confidence interval; HR, hazard ratio; mUICC, modified Union for International Cancer Control.

* Missing data was imputed.

† Crude hazard ratio for multicenter vs. single-center cohort.

‡ Curative treatment was defined as surgical resection, liver transplantation, and local ablation therapy.

§ Non-curative treatment was defined as TACE/TARE, radiotherapy, and systemic therapy.

Discussion

In this outcome-comparison study, we found that the single-center cohort (AMC group) was generally associated with significantly higher overall survival than the multicenter cohort (KPLCR group), and that this was consistent across all initial treatment modalities except systemic therapy and best supportive care.

These findings are noteworthy because to the best of our knowledge, this is the first retrospective cohort study to compare the overall survival of all-staged HCC patients in two large cohorts, one a nationwide multicenter cohort, the other a single-center cohort of representative volume. The retrospective design reflects real-life clinical practice in HCC patients with heterogeneous tumor features and variable prognoses, whereas this may be limited in RCTs as they involve highly-selected patient populations enrolled under strict eligibility criteria.[5] The differences observed between the two cohorts in overall survival are consistent with the findings of past studies that have compared the treatment outcomes of single-center and multicenter RCTs. These earlier studies showed that single-center RCTs produced larger treatment effects than multicenter RCTs,[9, 10, 23] and a review article has also highlighted the limited external validity of single-center RCTs by noting many instances in intensive care medicine in which the positive treatment outcomes found in single-center studies were not confirmed in multicenter RCTs.[8] However, the validity of retrospective studies of single-center cohorts has not been examined despite its importance.

The better overall survival observed above for systemic therapy and best supportive care in the multicenter cohort compared to the single-center cohort may be attributed to a center effect: in a previous study, patients who visited tertiary hospitals tended to receive more chemotherapy than patients who visited hospitals of secondary or primary levels.[24] In the tertiary hospital chosen as the single center in our investigation, a greater proportion of patients with unpreserved liver function received systemic treatment or best supportive care than in the multicenter series. As the survival of HCC patients is primarily dependent on baseline liver function,[5] one might anticipate that clinical outcomes would be poorer in the single-center cohort in patients with on average poorer liver function receiving systemic therapy and best supportive care.

On the other hand, the association of the single-center cohort with better survival outcomes for both surgical and loco-regional treatment modalities is likely to be related to the use of relatively homogeneous indications and the provision of standardized interventions by teams of high expertise in

high volume single-centers.[10, 23, 25] In addition, treatment outcomes obtained at different centers with varying treatment strategies and levels of experience, especially for difficult-to-treat cases, may not directly reflect the setting of any particular center-favorable outcomes in large centers may be overshadowed by the inclusion of a number of small volume centers with higher mortality in the multicenter series.[26] This may apply especially to HCC, as patients of the same stage can be treated differently due to individual tumor features as well as the variety of available or feasible treatment modalities, specific indications, and levels of skill and expertise in the different healthcare centers.[5]

Because of this heterogeneity, we established sub-cohorts to additionally compare the survival outcomes of treatments administered strictly according to the BCLC algorithm and the Milan criteria. These gave variable results; while there were no differences in overall survival between the two sub-cohorts for surgically and systemically-treated patients with favorable liver function, and transplant patients of any level of function, the multicenter cohort was associated with a significantly higher risk of mortality in patients who were locally treated with TACE or LAT as a standard option. The absence of a difference between patients who received liver transplants may be explained by the evidence that postoperative survival is not associated with transplant center volume, but is more likely attributable to other factors including donor age and patient characteristics such as age and MELD score.[27] Similarly, there was no significant difference in overall survival following surgical resection among Child-Pugh class A patients, as in studies that found no association between center type or volume and overall survival after surgical treatment of various cancers.[28-30] Surgical resection in most cases results in complete removal of the neoplasm,[31] making it an effective choice of curative treatment in patients who satisfy the indications. Also, advances in surgical technique and perioperative management may have decreased the gap in treatment outcomes between centers, at least for cases with preserved function.[32] Survival outcomes of systemic therapy also did not differ between the two sub-cohorts with good hepatic function as opposed to other malignancies.[33, 34] The lack of difference in survival outcomes for systemic therapy was probably related to the period when the study was performed: until 2018, sorafenib was the only approved treatment option for advanced HCC and it had only a modest survival benefit.[35] As numerous anticancer drugs for HCC have been approved since 2018,[36, 37] we believe that further studies are required to examine this interpretation.

The survival outcomes of TACE and LAT were, however, significantly different in the Child-Pugh class A subset: the multicenter cohort was associated with a higher risk of mortality than the single-center cohort, as observed in the complete cohorts. This finding may be attributable to the specialized nature

of these modalities and hence the influence that the interventional radiologists' skill and experience have upon the risk of recurrence as well as on post-procedural morbidity and mortality.[38-40] In this context, previous studies have shown that differences in skill have a greater impact on the efficacy of non-pharmacologic interventions than pharmacologic ones, as the level of expertise of care providers plays a more significant role in the former.[41-43] This may also explain why we detected significant differences in overall survival between the two cohorts in patients who received TACE or LAT, but not in those who received systemic therapy.

This study has potential limitations, which are mostly inherent in the retrospective nature of the study and the nature of the corresponding data sources. The variables reported, especially in the nationwide data, lacked some details such as family history of cancer, smoking status, and specific grade of performance. Additionally, data on disease recurrence and specific cause of death were unavailable and, as a result, the impact of disease recurrence and subsequent treatment on overall survival could not be assessed. Because recurrence or progression is common in HCC, progression-free survival might provide additional information regarding comparative treatment outcomes.[44] Completeness of the datasets was another issue, but we treated unavailable data in two ways to deal with that issue. We included the results of analyses performed with missing data classified both as a category and with the missing data substituted by multiple imputation, and we showed that the results obtained with the two methods did not differ significantly. Another possible limitation may be selection bias. The single-center cohort included a significantly higher proportion of early-stage patients according to BCLC staging (BCLC stage 0 or A) than the multicenter cohort and consequently the frequency of curative treatment as initial modality was higher, and the frequency of best supportive care lower in the single-center cohort than in the multicenter cohort. Because the data was retrospectively collected from patients randomly selected among various institutions, the KPLCR database did not provide per-sample information on the proportion and type of hospitals comprising the registry; however, patients were sampled using the probability proportional to size method and by regional stratification, which suggests that there would have been a balanced distribution of center volumes.[16] Furthermore, the results are probably reliable, as there were no significant differences between the results of univariate analysis and multivariable analysis after adjustment for established prognostic variables such as tumor stage and liver function. Lastly, the single-center data in our series were recruited from the highest-volume hospital in South Korea, and this could have led to the superior outcomes in terms of several modalities compared to the multicenter data. In general, however, the amount of retrospective HCC data from a

low-volume single-center would not be sufficient to provide less bias and adequate statistical power, and so would undermine the purpose of this study.

Conclusion

Comparison of overall survival between multicenter and single-center cohorts of patients with HCC showed significant differences in long-term outcomes according to primary treatment modality, but the differences were minimal in patients who received surgical resection, liver transplantation, or systemic therapy limited to the BCLC-guided treatments. The prognostic discrepancies between the two retrospective cohorts suggest that retrospective single-center studies should be interpreted with caution, particularly when evaluating HCC treatment outcomes beyond the BCLC criteria, and should involve careful consideration of center volume and study population. In short, good generalizability of treatment outcomes may still require collaboration between multiple centers.

References

[1] Velengtas P, Mohr P, Messner DA. Making informed decisions: assessing the strengths and weaknesses of study designs and analytic methods for comparative effectiveness research. National Pharmaceutical Council 2012.

[2] Galle PR, Forner A, Llovet JM, Mazzaferro V, Piscaglia F, Raoul J-L, et al. EASL clinical practice guidelines: management of hepatocellular carcinoma. Journal of hepatology 2018;69:182-236.

[3] Heimbach JK, Kulik LM, Finn RS, Sirlin CB, Abecassis MM, Roberts LR, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology 2018;67:358-380.

[4] 2022 KLCA-NCC Korea practice guidelines for the management of hepatocellular carcinoma. Clin Mol Hepatol 2022;28:583-705.

[5] Park JW, Chen M, Colombo M, Roberts LR, Schwartz M, Chen PJ, et al. Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. Liver International 2015;35:2155-2166.

[6] Euser AM, Zoccali C, Jager KJ, Dekker FW. Cohort studies: prospective versus retrospective. Nephron Clinical Practice 2009;113:c214-c217.

[7] Meinart C. Single-center versus multicenter trials. Clinical Trials: Design, Conduct and Analysis Meinart CL, Tonascia S (Eds) New York, Oxford University Press 1986:23-30.

[8] Bellomo R, Warrillow SJ, Reade MC. Why we should be wary of single-center trials. Critical care medicine 2009;37:3114-3119.

[9] Unverzagt S, Prondzinsky R, Peinemann F. Single-center trials tend to provide larger treatment effects than multicenter trials: a systematic review. Journal of clinical epidemiology 2013;66:1271-1280.

[10] Dechartres A, Boutron I, Trinquart L, Charles P, Ravaud P. Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. Annals of internal medicine 2011;155:39-51.

[11] Drakulovic MB, Torres A, Bauer TT, Nicolas JM, Nogué S, Ferrer M. Supine body position as a risk factor for nosocomial pneumonia in mechanically ventilated patients: a randomised trial. The Lancet 1999;354:1851-1858.

[12] van Nieuwenhoven CA, Vandenbroucke-Grauls C, van Tiel FH, Joore HC, van Schijndel RJS, van der Tweel I, et al. Feasibility and effects of the semirecumbent position to prevent ventilator-associated pneumonia: a randomized study. Critical care medicine 2006;34:396-402.

[13] Kim H-S, Lee S, Kim JH. Real-world evidence versus randomized controlled trial: clinical research based on electronic medical records. Journal of Korean medical science 2018;33.

[14] Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. Journal of biomedical informatics 2019;95:103208.

[15] Jin S, Choi W-M, Shim JH, Lee D, Kim KM, Lim Y-S, et al. Subclassification of advanced-stage hepatocellular carcinoma with macrovascular invasion: combined transarterial chemoembolization and radiotherapy as an alternative first-line treatment. Journal of Liver Cancer 2023;23:177-188.

[16] Chon YE, Lee HA, Yoon JS, Park JY, Kim BH, Lee IJ, et al. Hepatocellular carcinoma in Korea between 2012 and 2014: an analysis of data from the Korean nationwide cancer registry. Journal of Liver Cancer 2020;20:135-147.

[17] Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. Journal of hepatology 2022;76:681-693.

[18] Kudo M, Kitano M, Sakurai T, Nishida N. General rules for the clinical and pathological study of primary liver cancer, nationwide follow-up survey and clinical practice guidelines: the outstanding achievements of the liver cancer study group of Japan. Digestive diseases 2015;33:765-770.

[19] Cho Y, Choi JW, Kwon H, Kim KY, Lee BC, Chu HH, et al. Transarterial chemoembolization for hepatocellular carcinoma: 2023 Expert consensus-based practical recommendations of the Korean Liver Cancer Association. Korean journal of radiology 2023;24:606.

[20] Gavriilidis P, Roberts KJ, Askari A, Sutcliffe RP, Liu P-H, Hidalgo E, et al. Evaluation of the current guidelines for resection of hepatocellular carcinoma using the Appraisal of Guidelines for Research and Evaluation II instrument. Journal of hepatology 2017;67:991-998.

[21] Cho Y, Kim BH, Park J-W. Overview of Asian clinical practice guidelines for the management of hepatocellular carcinoma: An Asian perspective comparison. Clinical and Molecular Hepatology 2023;29:252.

[22] Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.

[23] Bafeta A, Dechartres A, Trinquart L, Yavchitz A, Boutron I, Ravaud P. Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. Bmj 2012;344.

[24] Rochigneux P, Raoul JL, Beaussant Y, Aubry R, Goldwasser F, Tournigand C, et al. Use of chemotherapy near the end of life: what factors matter? Ann Oncol 2017;28:809-817.

[25] Samaga D, Hornung R, Braselmann H, Hess J, Zitzelsberger H, Belka C, et al. Single-center versus multi-center data sets for molecular prognostic modeling: a simulation study. Radiation Oncology 2020;15:1-14.

[26] McCarthy JF, McCarthy PM, Massad MG, Cook DJ, Smedira NG, Kasirajan V, et al. Risk factors for death after heart transplantation: does a single-center experience correlate with multicenter registries? The Annals of thoracic surgery 1998;65:1574-1578.

[27] Northup P, Pruett T, Stukenborg G, Berg C. Survival after adult liver transplantation does not correlate with transplant center case volume in the MELD era. American journal of transplantation 2006;6:2455-2462.

[28] Berger NG, Silva JP, Mogal H, Clarke CN, Bedi M, Charlson J, et al. Overall survival after resection of retroperitoneal sarcoma at academic cancer centers versus community cancer centers: an analysis of the National Cancer Data Base. Surgery 2018;163:318-323.

[29] Meyerhardt JA, Tepper JE, Niedzwiecki D, Hollis DR, Schrag D, Ayanian JZ, et al. Impact of hospital procedure volume on surgical operation and long-term outcomes in high-risk curatively resected rectal cancer: findings from the Intergroup 0114 Study. Journal of clinical oncology 2004;22:166-174.

[30] Kozower BD, Stukenborg GJ. Hospital esophageal cancer resection volume does not predict patient mortality risk. The Annals of thoracic surgery 2012;93:1690-1698.

[31] Ferrari FS, Stella A, Pasquinucci P, Vigni F, Civeli L, Pieraccini M, et al. Treatment of small hepatocellular carcinoma: a comparison of techniques and long-term results. European journal of gastroenterology & hepatology 2006;18:659-672.

[32] Do You D, Kim DG, Seo CH, Choi HJ, Yoo YK, Park YG. Prognostic factors after curative resection hepatocellular carcinoma and the surgeon's role. Annals of Surgical Treatment and Research 2017;93:252-259.

[33] Haj Mohammad N, Bernards N, Besselink MG, Busch OR, Wilmink JW, Creemers GJ, et al. Volume matters in the systemic treatment of metastatic pancreatic cancer: a population-based study in the Netherlands. J Cancer Res Clin Oncol 2016;142:1353-1360.

[34] Malik AT, Alexander JH, Khan SN, Scharschmidt TJ. Is Treatment at a High-volume Center Associated with an Improved Survival for Primary Malignant Bone Tumors? Clin Orthop Relat Res 2020;478:631-642.

[35] Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc J-F, et al. Sorafenib in Advanced Hepatocellular Carcinoma. New England Journal of Medicine 2008;359:378-390.

[36] Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet 2017;389:56-66.

[37] Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim T-Y, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. New England Journal of Medicine 2020;382:1894-1905.

[38] Lee T-Y, Lin J-T, Ho HJ, Wu M-S, Wu C-Y. Evaluation of the effect of cumulative operator experience on hepatocellular carcinoma recurrence after primary treatment with radiofrequency ablation. Radiology 2015;276:294-301.

[39] Poon RT, Ng KK, Lam CM, Ai V, Yuen J, Fan ST, et al. Learning curve for radiofrequency ablation of liver tumors: prospective analysis of initial 100 patients in a tertiary institution. Annals of surgery 2004;239:441.

[40] Raoul J-L, Sangro B, Forner A, Mazzaferro V, Piscaglia F, Bolondi L, et al. Evolving strategies for the management of intermediate-stage hepatocellular carcinoma: available evidence and expert opinion on the use of transarterial chemoembolization. Cancer treatment reviews 2011;37:212-220.

[41] Biau DJ, Halm JA, Ahmadieh H, Capello WN, Jeekel J, Boutron I, et al. Provider and center effect in multicenter randomized controlled trials of surgical specialties: an analysis on patient-level data. Annals of surgery 2008;247:892-898.

[42] Boutron I, Tubach F, Giraudeau B, Ravaud P. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. Jama 2003;290:1062-1070.

[43] Jo D. The interpretation bias and trap of multicenter clinical research. The Korean Journal of Pain 2020;33:199-200.

[44] Llovet JM, Villanueva A, Marrero JA, Schwartz M, Meyer T, Galle PR, et al. Trial design and endpoints in hepatocellular carcinoma: AASLD consensus conference. Hepatology 2021;73:158-191.

Abstract

Background & Aims

We aimed to evaluate the validity of retrospective single-center versus multicenter research by comparing overall survival (OS) after various treatments in a nationwide multicenter cohort of hepatocellular carcinoma (HCC) patients with OS in a single-center cohort.

Methods

Patients newly diagnosed with HCC between January 2008 and December 2018 were analyzed using data from the Korean Primary Liver Cancer Registry (multicenter cohort, n=16,443), and the Asan Medical Center HCC registry (single-center cohort, n=15,655). Primary outcome was OS after initial treatment, which was compared between the two cohorts for both the entire population and for sub-cohorts with Child-Pugh A liver function (n=2,797 and n=5,151, respectively) treated according to the Barcelona-Clinic-Liver-Cancer (BCLC) strategy.

Results

Patients of BCLC stages 0 and A (59.3% vs. 35.2%) and patients who received curative treatment (42.1% vs. 32.1%) were more frequently observed in the single-center cohort (*Ps*<0.001). Multivariable analysis revealed worse OS in the multicenter cohort in patients receiving curative (adjusted hazard ratio [95% confidence interval], 1.48 [1.39–1.59]) and non-curative (1.22 [1.17–1.27]) treatments, and better OS in those receiving systemic therapy (0.83 [0.74–0.92]) and best supportive care (0.85 [0.79–0.91]). Subcohort analyses revealed differences in OS between the two cohorts in the subgroups undergoing chemoembolization (1.72 [1.48–2.00]) and ablation (1.44 [1.08–1.92]), with poorer OS in the multicenter sub-cohort.

Conclusions

Comparisons of treatment outcomes between single-center and multicenter cohorts revealed significant differences. Therefore, the results of retrospective single-center cohort studies of HCC treatments may not be generalizable to real-world practice.