Original Article

A machine learning model for colorectal liver metastasis post-hepatectomy prognostications

Cynthia Sin Nga Lam[1#], Alina Ashok Bharwani[1#], Evelyn Hui Yi Chan[1*], Vernice Hui Yan Chan[1*], Howard Lai Ho Au[1], Margaret Kay Ho[1**], Shireen Rashed[1**], Bernard Ming Hong Kwong[1], Wentao Fang[1], Ka Wing Ma[2], Chung Mau Lo[2], Tan To Cheung[2]**

[1]Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China; [2]Department of Surgery, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

Contributions: (I) Conception and design: CSN Lam, AA Bharwani, KW Ma, TT Cheung; (II) Administrative support: TT Cheung; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: AA Bharwani, EHY Chan, VHY Chan, HLH Au, MK Ho, S Rashed, BMH Kwong, W Fang; (V) Data analysis and interpretation: CSN Lam, AA Bharwani; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work as co-first authors.

[*]These authors contributed equally to this work.

[**]These authors contributed equally to this work.

Correspondence to: Dr. Tan To Cheung. Department of Surgery, Queen Mary Hospital, Pok Fu Lam Rd, Pokfulam, Hong Kong SAR, China. Email: cheung68@hku.hk.

Background: Currently, surgical resection is the mainstay for colorectal liver metastases (CRLM) management and the only potentially curative treatment modality. Prognostication tools can support patient selection for surgical resection to maximize therapeutic benefit. This study aimed to develop a survival prediction model using machine learning based on a multicenter patient sample in Hong Kong.

Methods: Patients who underwent hepatectomy for CRLM between 1 January 2009 and 31 December 2018 in four hospitals in Hong Kong were included in the study. Survival analysis was performed using Cox proportional hazards (CPH). A stepwise selection on Cox multivariable models with Least Absolute Shrinkage and Selection Operator (LASSO) regression was applied to a multiply-imputed dataset to build a prediction model. The model was validated in the validation set, and its performance was compared with that of Fong Clinical Risk Score (CRS) using concordance index.

Results: A total of 572 patients were included with a median follow-up of 3.6 years. The full models for overall survival (OS) and recurrence-free survival (RFS) consist of the same 8 established and novel variables, namely colorectal cancer nodal stage, CRLM neoadjuvant treatment, Charlson Comorbidity Score, pre-hepatectomy bilirubin and carcinoembryonic antigen (CEA) levels, CRLM largest tumor diameter, extrahepatic metastasis detected on positron emission-tomography (PET)-scan as well as KRAS status. Our CRLM Machine-learning Algorithm Prognostication model (CMAP) demonstrated better ability to predict OS (C-index =0.651), compared with the Fong CRS for 1-year (C-index =0.571) and 5-year OS (C-index =0.574). It also achieved a C-index of 0.651 for RFS.

Conclusions: We present a promising machine learning algorithm to individualize prognostications for patients following resection of CRLM with good discriminative ability.

Keywords: Machine-learning; colorectal liver metastasis (CRLM); prognostic model; survival analysis; hepatectomy outcome

Submitted Oct 31, 2021. Accepted for publication Feb 25, 2022. Published online Jul 12, 2022.
doi: 10.21037/hbsn-21-453
View this article at: https://dx.doi.org/10.21037/hbsn-21-453

Introduction

Colorectal cancer (CRC) has the third highest incidence and second highest mortality among all malignancies globally (1). In Hong Kong, CRC is the most common malignancy and second highest cause of cancer-related mortality, with over 20% of patients having metastasis at initial presentation (2). The liver is the most common site, accounting for approximately 50% of colorectal cancer metastases (3).

Despite contemporary advances in chemotherapy, hepatectomy remains the only curative treatment option for resectable CRLM (4). However, recurrence occurs in nearly three-quarters of patients within 16 months post-resection (5). As such, prognostication tools have been developed for the selection of patients to receive surgical resection, to maximize therapeutic benefit (Table S1).

Of these, the most commonly cited is the Fong Clinical Risk Score (CRS) for Colorectal Cancer Recurrence (6), which considers overall survival (OS) based on 5 variables. As all factors were equally weighted in a multivariable analysis, this may have oversimplified the prognostic impact of each. Since it was based on a single institution with a limited number of patients, its external validity and applicability in diverse population groups remain controversial (7,8). Aside from CRS, other older but well-established models include those from Nordlinger *et al.* (9) and Iwatsuki *et al.* (10).

More recent scoring systems such as the Genetic And Morphological Evaluation (GAME) score (11) and modified clinical score (m-CS) (12) have adapted to the modern age of genomic and chemotherapeutic advancements by including KRAS mutational status. However, as scores, they still suffer from the same limitations as CRS and are not as widely validated. There have been fewer studies on Asia-Pacific populations, with existing literature based on single-institution samples in Korea and China respectively (13,14) and a multi-center cohort in Japan (15).

This study aimed to develop a survival prediction model using machine learning on a multicenter patient sample in Hong Kong incorporating established and novel demographic, and clinicopathological variables. We present this article in accordance with the TRIPOD reporting checklist (available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-453/rc).

Methods

Ethical approval

All research processes were conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Institutional Review Board of The University of Hong Kong and the Hong Kong West Cluster of Hospital Authority, Hong Kong (Reference number UW 21-471).

Data source and variables

Data were anonymously extracted from electronic health records from the Clinical Management System (CMS) and Clinical Data Analysis and Reporting System (CDARS), an electronic healthcare database managed by the Hong Kong Hospital Authority. Patient-related variables were extracted from patient records. This included dates of CRC and CRLM diagnoses, dates of colorectal resection and hepatectomy, sex, age, Charlson Comorbidity Score (16), pre-hepatectomy bilirubin and carcinoembryonic antigen (CEA) levels, and the use of systemic treatments including neoadjuvant and adjuvant chemotherapy/radiotherapy for CRC and CRLM. Post-hepatectomy variables including follow-up status at the conclusion of the study, date of last follow-up, date of death, postoperative mortality and cause of death were also recorded if applicable.

Surgical variables for CRC [e.g., primary tumor location, operation category (emergency or elective)], hepatectomy approach (laparoscopic or open), hepatectomy extent [major involving 3 or more segments, or minor operation involving less than 3 segments (17)] and whether patients underwent combined radiofrequency ablation (RFA) treatment were recorded from operation reports. Pathological variables for CRC, namely the largest tumor diameter, lymphovascular invasion, venous infiltration, resection margin, tumor differentiation, number of positive lymph nodes, and KRAS status; for CRLM, namely number of tumor nodules, largest tumor diameter, lymphovascular invasion, resection margin, lobar involvement, and the new Edmondson grading for differentiation were extracted from pathology reports.

Findings from the positron emission-tomography (PET) imaging reports for CRLM were also noted, including tumor location, number of tumors, largest tumor diameter, and the presence of extrahepatic metastasis along with the

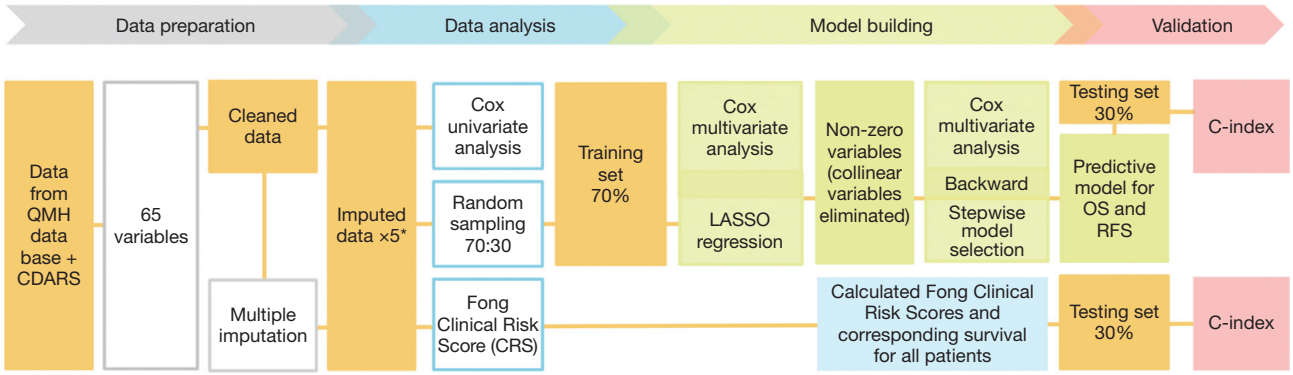

Figure 1 The research pipeline for data preparation, analysis as well as model building and validation. All results from analyses on imputed datasets were pooled. QMH, Queen Mary Hospital; CDARS, Clinical Data Analysis and Reporting System; LASSO, Least Absolute Shrinkage and Selection Operator; OS, overall survival; RFS, recurrence-free survival.

number and location of extrahepatic lesions. Additionally, whether extrahepatic lesions were resected and the timing of resection were also noted when applicable.

The manuscript adheres to the Transparent Reporting of multivariable prediction models for Individual Prognosis or Diagnosis (TRIPOD) statement (18). Investigators were not blinded from the outcome or predictor variables.

Study population

Patients who underwent hepatectomy for CRLM between 1 January 2009 and 31 December 2018 in four hospitals in Hong Kong (two teaching hospitals and two peripheral hospitals) were included in this study. Patients were excluded for the following: (I) unresected primary colorectal malignancy before/during the entire study period, (II) first liver metastasis treated by RFA only, (III) postoperative mortality, and (IV) multiple synchronous primary malignancies.

Statistical analysis

Statistical analysis and model building were performed using R 4.0.6. (*Figure 1*).

Data preparation

Using the mice (Multivariate Imputation by Chained Equations) statistical package (19), multiple imputation was performed to handle missing data and to increase statistical power and reduce bias. Missing values were replaced by five plausible data values, predicted upon other baseline values and/or outcomes. For all imputations, the results generated from analyses of the multiply imputed data were pooled, i.e., the results were averaged and the total variance was computed over the repeated analyses by Rubin's rules (20). Multiple imputation has been shown to provide more validity than simple ad-hoc approaches such as excluding entries with missing data or mean imputation (21,22). Certain continuous variables were made categorical using cutoffs derived from existing scoring systems including preoperative CEA level and bilirubin levels, whereas and the number of colorectal cancer lymph nodes affected was grouped according to TNM nodal status staging. OS (time between date of hepatectomy and the date of death), and recurrence-free survival (RFS) (time between the date of hepatectomy and the date of recurrence) were calculated.

Data analysis

Univariate Cox proportional hazards analysis was performed on the data using the survival package (23), investigating the association between all variables and the OS and RFS, as well as their statistical significance.

Model building and selection

Data was randomized into training and validation sets in a 70:30 ratio. To address multicollinearity, arising from the inevitable correlations between clinical predictors, Least Absolute Shrinkage and Selection Operator (LASSO) regression was performed using the glmnet package (24) to shrink the estimates of collinear predictors to zero and improve stability of predictions (25). With Cox multivariable analysis as the foundation of the model, backward model selection was applied on the non-zero

variables using the mice package (19). Each variable was removed in turn, and the pooled likelihood ratio P value was calculated and compared between the Cox multivariable models with and without the variable. If the P value of the comparison between the two models was >0.05, the corresponding variable would be removed. The procedure was repeated on the smaller model until all significant variables were removed.

The predictive model built on the training set predicts both OS and RFS, and was internally validated on the validation set, using Harrell's Concordance index (C-index), which handles censored data (26). The C-indices of our model were compared with those of CRS. Comparison with other more recent scoring systems such as GAME and m-CS could not be done since the predicted survival probabilities at different time points were not accessible.

Results

Baseline characteristics

Six hundred and fifty-two patients were screened, and a total of 572 patients were included with a median follow-up of 3.6 years [interquartile range (IQR): 2.5–5.7]. During the 9-year study period, 329 patients died. The important clinical and pathologic characteristics of the population are summarized in *Table 1*, and further detailed in Table S2. 62.9% of patients were male, and the median age was 62 (range, 24–94). Overall, left-sided CRC was the most common primary tumor location (38.5%). 61.4% patients presented with synchronous liver metastasis, but only 30.4% received synchronous resection. The median Charlson Comorbidity Index was 10 (IQR: 9.0–11.0). The median pre-CRC resection CEA was 9.7 ng/mL (IQR: 3.9–42.0), and the median pre-hepatectomy CEA was 12.0 ng/mL (IQR: 4.6–45.4). 33.3% of patients received 6 cycles of 5-FU-based neoadjuvant chemotherapy prior to hepatectomy.

Regarding pathology, the median largest CRC and CRLM diameters were 4.0 cm (IQR: 3.0–5.0) and 3.0 cm (IQR: 2.0–4.0) respectively. 41.2% of patients had KRAS mutations. More than two thirds of the patients (69.3%) had unilobar liver metastasis, and the median number of CRLM nodules was 2 (IQR: 1.0–3.0; range, 0–8). This result was consistent with the preoperative PET scan, which also showed 10.6% of patients having extrahepatic metastasis.

There were no statistically significant differences between the data distributions of the training and validation sets (Table S2).

Table 1 Baseline characteristics of the population

Variables	Total (n=572)
Sex, n (%)	
Male	360 (62.9)
Female	212 (37.1)
Age (years), median (Q1, Q3); range (min–max)	62 (56.8, 69.0); (24–94)
Charlson Comorbidity Score, median (Q1, Q3); range (min–max)	10 (9.0, 11.0); (8–16)
Presence of co-existing cancers (Hong Kong Top 10), n (%)	
No	557 (97.4)
Yes	15 (2.6)
Number of lymph nodes positive in primary CRC, median (Q1, Q3)	1 (0, 4.0)
Colorectal liver metastasis	
Pre-treatment PET scan findings[a]	
Number of tumor nodules, median (Q1, Q3); range (min–max)	1 (1.0, 3.0); (0–12)
Tumor location, n (%)	
Unilobar	343 (70.6)
Bilobar	143 (29.4)
Largest tumor diameter (cm), median (Q1, Q3); range (min–max)	2.5 (1.6, 4.0); (0–16.8)
Extrahepatic metastasis, n (%)	
No	423 (89.4)
Yes	50 (10.6)
Extrahepatic metastasis site, n (%)	
Lung metastasis	28 (56.0)
Bone metastasis	6 (12.0)
Other metastasis	16 (32.0)
Extrahepatic metastasis number, n (%)	
Single extrahepatic metastasis	47 (94.0)
Multiple extrahepatic metastases	3 (7.0)
Extrahepatic metastasis resection status, n (%)	
Not resected	28 (56.0)
Resected before hepatic resection	10 (20.0)
Resected during hepatic resection	6 (12.0)
Resected after hepatic resection	6 (12.0)

Table 1 (*continued*)

Table 1 (*continued*)

Variables	Total (n=572)
Pre-operative investigations	
Albumin (g/L), median (Q1, Q3)	40.0 (36.0, 43.0)
Bilirubin (μmol/L), median (Q1, Q3); range (min–max)	8.2 (6.0, 12.9); (2.0–51.0)
Liver CEA (ng/mL), median (Q1, Q3); range (min–max)	12.0 (4.6, 45.4); (0.7–4,040.0)
Liver pathology report findings	
Largest diameter of liver metastasis (cm), median (Q1, Q3); range (min–max)	3.0 (2.0, 4.0); (0.4–18.0)
Number of tumor nodules, median (Q1, Q3); range (min–max)	2 (1.0,3.0); (0–8)
Lymphovascular invasion, n (%)	
No	84 (52.5)
Yes	76 (47.5)
Tumor lobar involvement, n (%)	
Unilobar	395 (69.3)
Bilobar	175 (30.7)
KRAS, n (%)	
No mutation	250 (58.8)
Mutation	175 (41.2)
Treatment received	
Neoadjuvant treatment, n (%)	
No	378 (66.7)
Yes	189 (33.3)
Adjuvant chemotherapy, n (%)	
No	161 (28.1)
Yes	411 (71.9)
Patients who received only neoadjuvant therapy, n (%)	
No	523 (91.6)
Yes	48 (8.4)
Patients who received only adjuvant therapy, n (%)	
No	302 (53.2)
Yes	266 (46.8)
Patients who received both neoadjuvant and adjuvant chemotherapy, n (%)	
No	427 (75.2)
Yes	141 (24.8)

Table 1 (*continued*)

Variables	Total (n=572)
Follow-up	
Status, n (%)	
Censored (alive at the end of the study or was lost to follow up)	243 (42.5)
Dead	329 (57.5)
RFS, median (Q1, Q3), days	419.0 (186.5, 1,159.8)
OS, median (Q1, Q3), days	1,169.0 (771.5, 1,893.0)
DFI, median (Q1, Q3), days	0 (0, 218.0)

This table summarizes patient data on key clinically significant variables only. Additional data on all the 65 variables extracted from patient records is summarized in Table S2. [a], For all patients, data was extracted from their pretreatment/baseline PET scan, where available; for patients who received neoadjuvant chemotherapy, data was extracted from their initial PET scan prior to commencing any neoadjuvant treatment; for the 24 patients who only had PET scan data available after commencing treatment, this was not considered to be representative of their baseline characteristics and thus their PET scan data was not included in analysis. CRC, colorectal cancer; CEA, carcinoembryonic antigen; PET, positron emission-tomography; RFS, recurrence-free survival; OS, overall survival; DFI, disease free interval.

Univariate cox proportional hazard analysis

Univariate analysis revealed that OS and RFS were significantly associated with 27 and 22 variables respectively (P<0.05) (Table S3). Significant variables for both OS and RFS included CRLM clinical risk scores and their parameters, CRC and CRLM pathological factors, liver PET scan parameters, and treatment factors. Patient demographic factors were not associated with either OS or RFS.

Final model results

LASSO regression returned 17 non-zero variables for both OS and RFS (*Table 2*). Nine additional variables were eliminated using backward model selection. The final models for OS and RFS were conducted with the same list of shortlisted 8 predictors (*Table 2*).

Kaplan-Meier curves plotted on the model's predictions on the imputed validation cohort demonstrated good accuracy compared with the actual survival in the same

Table 2 Predictors considered and selected for the final models

	Overall survival		Recurrence-free survival	
Final model	Global P value: 2.583e-08		Global P value: 6.005e-06	
	Survival probability is 1 at baseline, which is the date of surgery		Survival probability is 1 at baseline, which is the date of surgery	
	Predictor variables	Pooled coefficient	Predictor variables	Pooled coefficient
	1. Bilirubin (>35 μmol/L) pre-hepatectomy	2.31	1. Bilirubin (>35 μmol/L) pre-hepatectomy	1.46
	2. Extrahepatic Metastasis detected on PET-scan	0.41	2. CEA pre-hepatectomy (>200 ng/mL)	0.51
	3. CEA pre-hepatectomy (>200 ng/mL)	0.25	3. Neoadjuvant treatment for CRLM	0.43
	4. Neoadjuvant treatment for CRLM	0.28	4. Extrahepatic Metastasis detected on PET-scan	0.30
	5. CRC N stage	0.25	5. CRC N stage	0.30
	6. KRAS mutation status	0.18	6. KRAS mutation status	0.18
	7. Charlson Comorbidity Score	0.17	7. Charlson Comorbidity Score	0.12
	8. CRLM largest tumor diameter	0.06	8. CRLM largest tumor diameter	0.06
Other non-0 predictors from LASSO regression excluded from the final model	1. Adjuvant Chemotherapy following CRLM resection		1. Adjuvant Chemotherapy following CRLM resection	
	2. Number of CRLM		2. Number of CRLM	
	3. CRLM lymphovascular invasion		3. Number CRLM >3 (Nordlinger)	
	4. Largest CRLM tumor diameter on PET-scan		4. CRLM lymphovascular invasion	
	5. CRC # of lymph nodes		5. Bilobar CRLM on PET-scan	
	6. CRC lymphovascular invasion		6. CRC lymphovascular invasion	
	7. CRC resection margin		7. CRC resection margin	
	8. Disease free interval >1 year (Fong Score)		8. Disease Free interval >1 year (Fong Score)	
	9. Laparoscopic hepatectomy		9. Laparoscopic hepatectomy	

LASSO, Least Absolute Shrinkage and Selection Operator; PET, positron emission-tomography; CEA, carcinoembryonic antigen; CRLM, colorectal liver metastases; CRC, colorectal cancer.

dataset, for both OS and RFS (*Figure 2*).

Our model gives OS and RFS predictions at continuous time points with a global C-index for OS and RFS respectively. Since CRS only gives 1-year and 5-year OS predictions, our C-indices could only be compared with those of CRS at these two timepoints. Validated in the imputed validation cohort, our model demonstrated better ability to predict OS, with a C-index of 0.651, compared with the CRS for 1-year and 5-year OS, with a relative improvement of 13.9% and 13.5% respectively. Our model achieved a C-index of 0.651 for RFS (*Table 3*).

Discussion

Given that long-term outcomes post-hepatectomy may vary widely among patients, prognostic tools are crucial in facilitating individualized patient care and multidisciplinary discussion and in identifying patients who may benefit from specific therapeutic modalities. While a few traditional scoring systems (6,9,10) have been validated in international cohorts, they largely rely on histopathological variables and do not consider newer treatment regimens or molecular variables. More recent scores (11,12) which address KRAS mutation status and neoadjuvant chemotherapy are not yet widely validated in external cohorts, and their integer risk scores inherently limit their accuracy in reflecting nonlinear and complex interaction effects between predictors, especially in the modern era of computed recording and prediction.

We developed and internally validated a model that predicts survival of patients with CRLM post-hepatectomy for colorectal liver metastasis using data from 572 patients

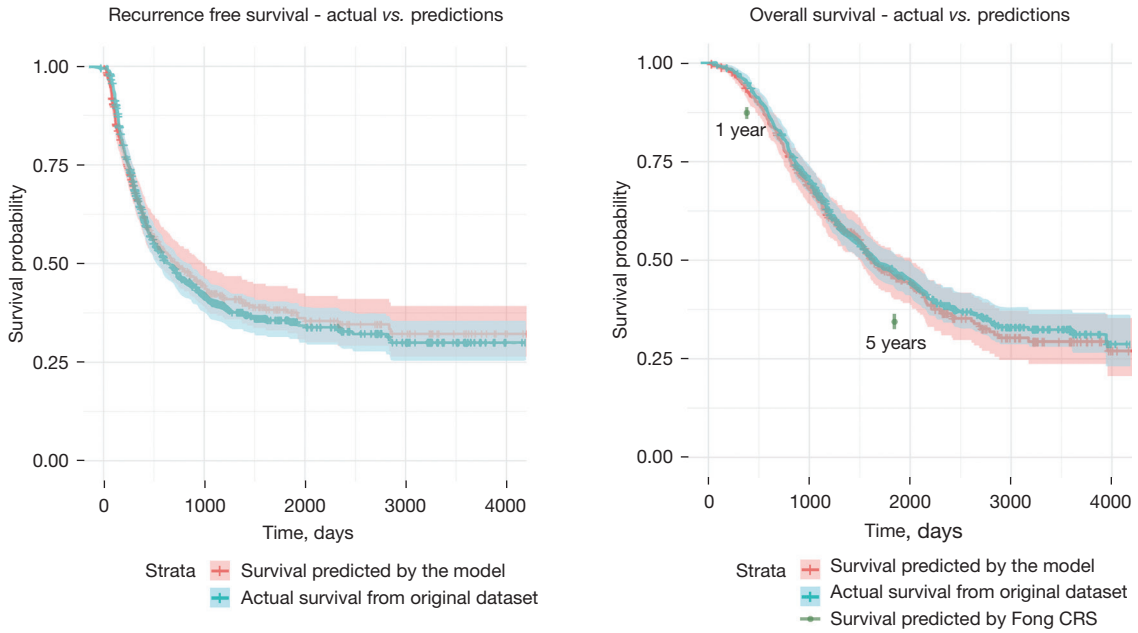

Figure 2 Kaplan-Meier curves plotted on the model's predictions on the imputed validation cohort, compared with the actual survival and the OS predicted by CRS. CRS, Clinical Risk Score; OS, overall survival.

Table 3 The performance of our model compared with CRS

	Our model	Fong CRS	Relative improvement
Overall survival			
1-year survival	0.651 (95% CI: 0.537–0.765)	0.571 (95% CI: 0.514–0.628)	13.9%
5-year survival		0.574 (95% CI: 0.517–0.630)	13.5%
Recurrence-free survival	0.651 (95% CI: 0.597–0.705)		

CRS, Clinical Risk Score; CI, confidence interval.

across 4 major public hospitals in Hong Kong. This is the first prognostication tool for CRLM patients in this locality. When validated in an internal validation set, our model CMAP demonstrated better ability to predict OS (C-index =0.651), compared with the Fong CRS for 1-year (C-index =0.571) and 5-year OS (C-index =0.574). It also achieved a C-index of 0.651 for RFS.

In addition to tailoring to the local population, CMAP accounts for recent advances in CRLM treatment by integrating holistic prognostic information about a patient's premorbid status, genetic profile, imaging and blood test findings as well as management in predicting patients' recurrence and survival post-hepatectomy. The need for individualized prognostication is being increasingly emphasized in the modern era of precision medicine (27). Instead of giving generalized estimates of survival by categorizing patients into subgroups, our model is able to account for 8 unique variables that are clinically important in the current standard of care and provide individualized predictions for recurrence and survival at continuous points of time. This can facilitate more discussions tailored to each patient's unique situation depending on their priorities and expectations of the disease course.

Furthermore, each variable of our model is independently weighed in our machine learning model to enable more accurate prediction of both OS and RFS. The a-CS model developed by Paredes *et al.* (28), like ours, incorporated machine learning to improve their robustness. Although they reported higher discriminatory ability than the CRS and m-CS, this comparison's significance is debatable as

their model calculates RFS, whereas CRS and m-CS predict OS (29). In our model-building process, predictors for RFS and OS were analyzed separately and while the final model consists of the same list of variables, each model accounts for different weights and interactions between the variables. To ensure a fair comparison, only the OS model's performance was compared with that of CRS, which only predicts OS.

Variables included in our model which are supported in existing scoring systems include preoperative CEA (6,11,14,15,28,30) size of liver metastases (6,9,10,12,14, 15,28,30), the presence of extrahepatic metastasis (11,14,15,28,30), and positive nodal status of primary CRC (6,11,12,30). Preoperative CEA has been found in studies to impact OS and RFS as well as response to systemic therapy (6,31). The largest tumor diameter of CRLM was used as a continuous rather than a dichotomized variable as studies have shown that binary cut-off values for this parameter may not accurately represent its prognostic significance (32). Existing literature reports that positive primary tumor nodal status in patients with resectable CRLM predicts poor RFS (33). The presence of extrahepatic metastasis was previously considered an absolute contraindication to hepatectomy; however, this notion has been challenged with the advent of novel systemic therapeutic agents which have expanded the criteria for surgical resectability (34). Our model joins recent scoring systems which have similarly included extrahepatic disease as a poor prognostic factor for both OS and RFS (11,14,28,30). Further investigation into specific factors impacting survival outcomes for surgically fit patients with extrahepatic disease should be conducted with the aim of identifying clearer determinants of surgical futility.

In addition to these variables, the current model also includes newer prognostic variables including the use of neoadjuvant chemotherapy and KRAS mutation status, similar to the recent m-CS and GAME score (11,12). Patients with KRAS mutant status have poorer prognostic outcomes given their associated resistance to Epidermal Growth Factor Receptor (EGFR)-targeted therapies as well as more aggressive tumor behavior (31); this has been echoed in existing studies (35,36).

Neoadjuvant chemotherapy pre-hepatectomy was found to be associated with worse OS and RFS, a result also reported to be of interest in other studies (11,28,37,38). This association may be due to the heterogeneity of chemotherapy regimens used, interactions between underlying factors such as tumor molecular status (11), or the fact that patients receiving neoadjuvant chemotherapy often have worse disease status and poorer surgical prognosis to begin with (28). Upon additional retrospective analysis of our study population, patients who received neoadjuvant therapy were more likely to have had bilobar liver metastasis (45.2% *vs.* 23.3%, P<0.001) and multiple liver lesions (64.9% *vs.* 42.3%, P<0.001), when compared to those directly undergoing hepatectomy. Future research should assess the prognostic impact of patients' baseline disease status as well as specific therapeutic regimens including chemotherapeutic agents used, number of cycles, and objective parameters to measure treatment response.

Our inclusion of Charlson Comorbidity Score and pre-hepatectomy bilirubin level reflects our effort to include clinically available and practical predictors, especially in this era where there is increasing evidence for the use of parenchymal-sparing approaches for patient tolerability (39). A high Charlson comorbidity score was associated with higher short-term and long-term mortality in patients by Robertson *et al.* (40), and was also deemed useful in predicting lower OS in CRC patients over 75 years of age in Japan (41). Yang *et al.* (42) reported that pre-treatment direct bilirubin was an independent prognostic factor for OS in stage IV CRC patients. Ma *et al.* (43) found that a bilirubin level of >35 mmol/L was a significant predictor of poor post-hepatectomy outcomes for primary and secondary liver tumors. Additionally, we utilized preoperative PET-CT results to detect the presence of extrahepatic metastasis. Although some studies suggest that preoperative PET-CT is not related to improvements in OS or DFS in CRLM patients (44), PET-CT has demonstrated high sensitivity for detecting distant metastases (45). With PET-CT becoming a more popular preoperative staging tool, future studies may fare better should they incorporate the value of PET findings in prognostic models.

However, certain variables included in previous scoring systems were not found to be significant within this model, including the number of metastatic hepatic lesions (6,9,10,15,28,30) and a shorter disease-free interval (DFI) (6,9,10,28). For the number of hepatic lesions, Jang *et al.* (46) also reported no statistically significant difference in OS between patients with 1–2 CRLM nodules, compared to those with 3–8 nodules. The advent of modern treatment modalities may explain similar survival outcomes regardless of the number of liver metastases (12). Similarly, other

newer scores have also excluded DFI as a variable including the m-CS (12,47) which did not find DFI to significantly predict OS. Conversely, there is expanding literature regarding the worse prognosis of metastases detected after surgical resection (i.e., metachronous lesions) and adjuvant therapy for the primary tumor compared to those synchronously detected, potentially due to unfavorable biological characteristics such as chemotherapy-resistance (12,48).

Our tool takes a holistic view of the patient, integrating prognostic information about their premorbid status, genetic profile, imaging and blood test findings as well as management, allowing for more individualized risk estimates that can be applied to complex clinical scenarios. It will be deployed as a web-based or phone-based clinical decision tool for clinicians to easily input 8 clinical data points. Since our tool supports the predictions of OS and RFS at continuous points in time, this can facilitate more discussions tailored to each patient's unique situation depending on their priorities and expectations of the disease course.

Limitations

Retrospective data extraction meant that potentially important investigation results and reports, especially those from in the private sector, were not always accessible through electronic health records. Variables with a high proportion of missing data included BRAF status (98.4%), New Edmondson grading for colorectal cancer histology (79.7%), and the presence of CRLM lymphovascular invasion (72.0%). Additionally, certain molecular variables such as NRAS status were not available on patient records. Although the study included consideration of potential confounders such as age and sex, others, like socioeconomic status were not closely examined; this may warrant further analysis. Furthermore, the retrospective nature of this study may have introduced potential selection bias with respect to the variables examined. This was minimized by a priori inclusion of 68 clinically available raw variables for analysis to avoid selective reporting or inclusion.

Data was collected on several other prognostic variables that have been reported in existing literature such as BRAF mutation status and PET findings, however there was insufficient information from our database to draw meaningful conclusions. BRAF mutation status was found by recent studies and a meta-analysis to be strongly associated with worse prognosis (49,50). However, information on BRAF status was only available for 1.6% of patients in our database, thus precluding further conclusions from being drawn. Tumor pathological response to chemotherapy has also been associated with improved survival (51,52). While we collected data on the use of neoadjuvant chemotherapy, data unavailability prevented measurement of tumor response by either radiological (e.g., RECIST) or circulating tumor cells (e.g., CyCAR) criteria, both found to be independent prognostic factors for OS (53).

One of the variables in our model, the largest liver tumor diameter, was determined post-operatively using pathology reports. This is a common limitation faced by other prognostication tools (6,11,32,54). It does not necessarily affect the utility of the tool since CT imaging, as the mainstay for CRLM diagnosis, can be readily substituted as an accurate tool for determining the size of metastases (55). However, future studies should attempt to build models solely with pre-operative findings and compare them with those built with post-operative findings to account for CT's limitations as well as the differences between the initial CT scan and operation.

The statistical methodology's weaknesses should also be considered when interpreting the results. Every statistical model comes with certain assumptions in which the model's validity can be threatened if they are not met (56). A key assumption in the Cox model is the proportional hazard assumption, meaning the ratio of hazards (the output of a Cox model) is constant over time. Future research should include external validation in diverse cohorts as well as prospective studies both locally and internationally to assess accuracy. In addition, Margonis *et al.* (29) pointed out the inherent limitation of building "One size fits all" prediction models in heterogeneous populations. To account for effect modification by factors that potentially affect CRLM prognosis, patients should be stratified into more homogenous groups according to variables such as comorbid status, extrahepatic metastases, and biological markers to increase accuracy and applicability. Future research can use machine learning to better develop personalized predictions for these sub-stratified populations as identified through the preliminary findings of clinically significant variables in this study. Additionally, given the rapid progress in the field of newer oncological therapies for late-stage cancers, we believe our model provides a backbone for future analyses in this locality, allowing for the incorporation of newer therapeutic regimens in future models.

Conclusions

We presented a promising machine learning algorithm for survival prediction for patients following resection of CRLM. Our model was built on multicenter patient data in Hong Kong, considered new variables and addressed missing data and multicollinearity. Our model demonstrated superior performance than CRS for OS prediction, and an excellent concordance index for both RFS and OS predictions.

In the future, other machine learning algorithms should also be considered to further improve the model performance. The deployment of the model, its evaluation and external validation is reserved for a future publication as it merits more discussion than can be included here.

Acknowledgments

We would like to thank Ms. Chrissie Chiu and Mr. Kim Yuen (Department of Surgery, QMH) for their kind support and guidance in data collection.
Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-453/rc

Data Sharing Statement: Available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-453/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-453/coif). CML serves as the unpaid editorial board member of *Hepatobiliary Surgery and Nutrition*. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All research processes were conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Institutional Review Board of The University of Hong Kong and the Hong Kong West Cluster of Hospital Authority, Hong Kong (Reference number UW 21-471).

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Overview of Hong Kong cancer statistics of 2018. Hong Kong Cancer Registry. 2020. Available online: http://www3.ha.org.hk/cancereg
3. Filip S, Vymetalkova V, Petera J, et al. Distant Metastasis in Colorectal Cancer Patients—Do We Have New Predicting Clinicopathological and Molecular Biomarkers? A Comprehensive Review. Int J Mol Sci 2020;21:5255.
4. Xu F, Tang B, Jin TQ, et al. Current status of surgical treatment of colorectal liver metastases. World J Clin Cases 2018;6:716-34.
5. Tabchouri N, Gayet B, Okumura S, et al. Recurrence patterns after laparoscopic resection of colorectal liver metastases. Surg Endosc 2018;32:4788-97.
6. Fong Y, Fortner J, Sun RL, et al. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. Ann Surg 1999;230:309-18; discussion 318-21.
7. Beamish P, Lemke M, Li J, et al. Validation of clinical risk score for colorectal liver metastases resected in a contemporary multicenter cohort. HPB (Oxford) 2017;19:675-81.
8. Mann CD, Metcalfe MS, Leopardi LN, et al. The clinical risk score: emerging as a reliable preoperative prognostic index in hepatectomy for colorectal metastases. Arch Surg 2004;139:1168-72.
9. Nordlinger B, Guiguet M, Vaillant JC, et al. Surgical resection of colorectal carcinoma metastases to the liver: a prognostic scoring system to improve case selection, based on 1568 patients. Association Française de Chirurgie. Cancer 1996;77:1254-62.
10. Iwatsuki S, Dvorchik I, Madariaga JR, et al. Hepatic resection for metastatic colorectal adenocarcinoma: a proposal of a prognostic scoring system. J Am Coll Surg

1999;189:291-9.
11. Margonis GA, Sasaki K, Gholami S, et al. Genetic And Morphological Evaluation (GAME) score for patients with colorectal liver metastases. Br J Surg 2018;105:1210-20.
12. Brudvik KW, Jones RP, Giuliante F, et al. RAS Mutation Clinical Risk Score to Predict Survival After Resection of Colorectal Liver Metastases. Ann Surg 2019;269:120-6.
13. Kim WJ, Lim TW, Kang SH, et al. Development and validation of novel scoring system for the prediction of disease recurrence following resection of colorectal liver metastasis. Asian J Surg 2020;43:438-46.
14. Liang JY, Lin HC, Liu J, et al. A novel prognostic nomogram for colorectal cancer liver metastasis patients with recurrence after hepatectomy. Cancer Med 2021;10:1535-44.
15. Beppu T, Sakamoto Y, Hasegawa K, et al. A nomogram predicting disease-free survival in patients with colorectal liver metastases treated with hepatic resection: multicenter data collection as a Project Study for Hepatic Surgery of the Japanese Society of Hepato-Biliary-Pancreatic Surgery. J Hepatobiliary Pancreat Sci 2012;19:72-84.
16. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83.
17. Dahiya D, Wu TJ, Lee CF, et al. Minor versus major hepatic resection for small hepatocellular carcinoma (HCC) in cirrhotic patients: a 20-year experience. Surgery 2010;147:676-85.
18. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015;350:g7594.
19. van Buuren S, Groothuis-Oudshoorn K, Robitzsch A. Package 'mice': Multivariate Imputation by Chained Equations. CRAN Repos. 2019. Available online: https://cran.r-project.org/web/packages/mice/mice.pdf
20. Rubin DB. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 2004.
21. Laird NM. Missing data in longitudinal studies. Stat Med 1988;7:305-15.
22. McCleary L. Using multiple imputation for analysis of incomplete data in clinical research. Nurs Res 2002;51:339-43.
23. Therneau TM, Lumley T, Elizabeth A, Cynthia C, Therneau MTM. Package 'survival'. 2020. Available online: https://CRAN.R-project.org/package=survival
24. Friedman JH, Hastie T, Tibshirani R. glmnet: lasso and elastic-net regularized generalized linear models, 2010b. 2020:1.-5. Available online: https://cran.r-project.org/web/packages/glmnet/glmnet.pdf
25. Dormann CF, Elith J, Bacher S, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 2013;36:27-46.
26. Harrell Jr FE, Harrell Jr MFE. Package 'hmisc'. CRAN2018. 2019;2019:235-6. Available online: https://cran.r-project.org/web/packages/Hmisc/Hmisc.pdf
27. Mahar AL, Compton C, Halabi S, et al. Personalizing prognosis in colorectal cancer: A systematic review of the quality and nature of clinical prognostic tools for survival outcomes. J Surg Oncol 2017;116:969-82.
28. Paredes AZ, Hyer JM, Tsilimigras DI, et al. A Novel Machine-Learning Approach to Predict Recurrence After Resection of Colorectal Liver Metastases. Ann Surg Oncol 2020;27:5139-47.
29. Margonis GA, Andreatos N, Brennan MF. Predicting Survival in Colorectal Liver Metastasis: Time for New Approaches. Ann Surg Oncol 2020;27:4861-3.
30. Rees M, Tekkis PP, Welsh FK, et al. Evaluation of long-term survival after hepatic resection for metastatic colorectal cancer: a multifactorial model of 929 patients. Ann Surg 2008;247:125-35.
31. Spolverato G, Ejaz A, Azad N, et al. Surgery for colorectal liver metastases: The evolution of determining prognosis. World J Gastrointest Oncol 2013;5:207-21.
32. Sasaki K, Morioka D, Conci S, et al. The tumor burden score: a new "metro-ticket" prognostic tool for colorectal liver metastases based on tumor size and number of tumors. Ann Surg 2018;267:132-41.
33. Seeberg LT, Brunborg C, Waage A, et al. Survival Impact of Primary Tumor Lymph Node Status and Circulating Tumor Cells in Patients with Colorectal Liver Metastases. Ann Surg Oncol 2017;24:2113-21.
34. Pawlik TM, Schulick RD, Choti MA. Expanding criteria for resectability of colorectal liver metastases. Oncologist 2008;13:51-64.
35. Karagkounis G, Torbenson MS, Daniel HD, et al. Incidence and prognostic impact of KRAS and BRAF mutation in patients undergoing liver surgery for colorectal metastases. Cancer 2013;119:4137-44.
36. Nash GM, Gimbel M, Shia J, et al. KRAS mutation correlates with accelerated metastatic progression in patients with colorectal liver metastases. Ann Surg Oncol 2010;17:572-8.
37. Passot G, Denbo JW, Yamashita S, et al. Is hepatectomy justified for patients with RAS mutant colorectal liver metastases? An analysis of 524 patients undergoing

curative liver resection. Surgery 2017;161:332-40.
38. Ito H, Are C, Gonen M, et al. Effect of postoperative morbidity on long-term survival after hepatic resection for metastatic colorectal cancer. Ann Surg 2008;247:994-1002.
39. Evrard S, Poston G, Kissmeyer-Nielsen P, et al. Combined ablation and resection (CARe) as an effective parenchymal sparing treatment for extensive colorectal liver metastases. PLoS One 2014;9:e114404.
40. Robertson DJ, Stukel TA, Gottlieb DJ, et al. Survival after hepatic resection of colorectal cancer metastases: a national experience. Cancer 2009;115:752-9.
41. Tominaga T, Nonaka T, Takeshita H, et al. The Charlson Comorbidity Index as an Independent Prognostic Factor in Older Colorectal Cancer Patients. Indian J Surg 2018;80:54-60.
42. Yang L, Ge LY, Yu T, et al. The prognostic impact of serum bilirubin in stage IV colorectal cancer patients. J Clin Lab Anal 2018;32:e22272.
43. Ma KW, Cheung TT, She WH, et al. Risk prediction model for major complication after hepatectomy for malignant tumor - A validated scoring system from a university center. Surg Oncol 2017;26:446-52.
44. Daza JF, Solis NM, Parpia S, et al. A meta-analysis exploring the role of PET and PET-CT in the management of potentially resectable colorectal cancer liver metastases. Eur J Surg Oncol 2019;45:1341-8.
45. Grut H, Dueland S, Line PD, et al. The prognostic value of 18F-FDG PET/CT prior to liver transplantation for nonresectable colorectal liver metastases. Eur J Nucl Med Mol Imaging 2018;45:218-25.
46. Jang KU, Kim CW, Kim KH, et al. Prognostic Factors in Terms of the Number of Metastatic Nodules in Patients With Colorectal Cancer Liver Metastases. Ann Coloproctol 2016;32:92-100.
47. Höppener DJ, Nierop PMH, van Amerongen MJ, et al. The Disease-Free Interval Between Resection of Primary Colorectal Malignancy and the Detection of Hepatic Metastases Predicts Disease Recurrence But Not Overall Survival. Ann Surg Oncol 2019;26:2812-20.
48. Andreou A, Kopetz S, Maru DM, et al. Adjuvant chemotherapy with FOLFOX for primary colorectal cancer is associated with increased somatic gene mutations and inferior survival in patients undergoing hepatectomy for metachronous liver metastases. Ann Surg 2012;256:642-50.
49. Tsilimigras DI, Ntanasis-Stathopoulos I, Bagante F, et al. Clinical significance and prognostic relevance of KRAS, BRAF, PI3K and TP53 genetic mutation analysis for resectable and unresectable colorectal liver metastases: A systematic review of the current evidence. Surg Oncol 2018;27:280-8.
50. Margonis GA, Buettner S, Andreatos N, et al. Association of BRAF Mutations With Survival and Recurrence in Surgically Treated Patients With Metastatic Colorectal Liver Cancer. JAMA Surg 2018;153:e180996.
51. Primavesi F, Fadinger N, Biggel S, et al. Early response evaluation during preoperative chemotherapy for colorectal liver metastases: Combined size and morphology-based criteria predict pathological response and survival after resection. J Surg Oncol 2020;121:382-91.
52. Berardi G, De Man M, Laurent S, et al. Radiologic and pathologic response to neoadjuvant chemotherapy predicts survival in patients undergoing the liver-first approach for synchronous colorectal liver metastases. Eur J Surg Oncol 2018;44:1069-77.
53. Delgado-Ureña M, Ortega FG, de Miguel-Pérez D, et al. Circulating tumor cells criteria (CyCAR) versus standard RECIST criteria for treatment response assessment in metastatic colorectal cancer patients. J Transl Med 2018;16:251.
54. Byun SS, Heo TS, Choi JM, et al. Deep learning based prediction of prognosis in nonmetastatic clear cell renal cell carcinoma. Sci Rep 2021;11:1242.
55. Cantisani V, Grazhdani H, Fioravanti C, et al. Liver metastases: Contrast-enhanced ultrasound compared with computed tomography and magnetic resonance. World J Gastroenterol 2014;20:9998-10007.
56. Freedman DA. Statistical models: theory and practice. Cambridge University press; 2009.

Cite this article as: Lam CSN, Bharwani AA, Chan EHY, Chan VHY, Au HLH, Ho MK, Rashed S, Kwong BMH, Fang W, Ma KW, Lo CM, Cheung TT. A machine learning model for colorectal liver metastasis post-hepatectomy prognostications. HepatoBiliary Surg Nutr 2023;12(4):495-506. doi: 10.21037/hbsn-21-453

Supplementary

Table S1 Existing scoring systems for patients with colorectal liver metastasis

Scoring system	Predicted outcome	Location	Study size	Parameters
Nordlinger *et al.*, 1996	Overall survival	France	1,568	• Age ≥60 • Serosal invasion of the primary tumor • Lymph node positive primary disease • Disease-free interval <24 months • No. of liver metastases >3 • Largest liver metastasis ≥5 cm • Margin ≤1 cm
Fong *et al.*, 1999	Overall survival	USA	1,001	• DFI <12 months • No. of liver metastases >1 • Preoperative CEA >200 ng/mL • Largest liver metastasis >5 cm • Lymph node positive primary disease
Iwatsuki *et al.*, 1999	Overall survival and recurrence-free survival	USA	305	• No. of liver metastases >2 • Bilobar lesions • Time from treatment of primary to diagnosis of CRLM ≤30 months • Tumor size >8 cm
Rees *et al.*, 2008	Cancer-specific survival	UK	929	• No. of liver metastases >3 • Lymph node positive primary disease • Poorly differentiated primary tumor • Extrahepatic metastasis • Largest liver metastasis ≥5 cm • CEA >60 ng/mL • Margin involved
Beppu *et al.*, 2012	Recurrence-free survival	Japan	727	• Synchronous metastases • Lymph node positive primary disease • No. of liver metastases 2–4, ≥5 • Largest liver metastasis >5 cm • Extrahepatic metastasis • Preoperative CA19-9 >100 U/mL
Sasaki *et al.*, 2018	Overall survival	USA	604	• Maximum tumor size • No. of liver metastases
Margonis *et al.*, 2018	Overall survival	USA	502	• KRAS mutation • CRC nodal metastases • Tumor burden score 3–8, ≥9 • CEA ≥20 ng/mL • Extrahepatic disease
Brudvik *et al.*, 2019	Overall and recurrence-free survival	USA	564	• Lymph node positive primary disease • Largest liver metastasis >5 cm • RAS Mutation

This is not an exhaustive list of all existing scoring systems.

https://dx.doi.org/10.21037/hbsn-21-453

Table S2 Baseline characteristics of the study population and comparison between the training and validation sets

Variables	Total (n=572)	Training set (n=400)	Testing set (n=172)	P value	Missing data, n (%)
Baseline characteristics					
Hospital, n (%)				0.176	
QMH	230 (40.2)	158 (39.5)	72 (41.9)		
QEH	101 (17.7)	65 (16.3)	36 (20.9)		
TMH	121 (21.1)	84 (21.0)	37 (21.5)		
PWH	120 (21.0)	93 (23.3)	27 (15.7)		
Sex, n (%)				0.423	
Male	360 (62.9)	247 (61.8)	113 (65.7)		
Female	212 (37.1)	153 (38.3)	59 (34.3)		
Age, median (Q1, Q3); range (min–max)	62 (56.8, 69.0); (24–94)	62 (56.0, 68.0)	63 (57.0, 70.0)	0.308	
Smoking, n (%)				0.724	42 (7.3)
No	348 (65.7)	240 (65.2)	108 (66.6)		
Moderate	130 (24.5)	94 (25.5)	36 (22.2)		
Heavy	52 (9.8)	35 (9.5)	17 (10.5)		
Drinking, n (%)				0.747	56 (9.8)
No	369 (71.5)	260 (72.2)	109 (69.9)		
Moderate	115 (22.3)	77 (21.4)	38 (24.4)		
Heavy	32 (6.2)	23 (6.4)	9 (5.8)		
Charlson Comorbidity Score, Median (Q1, Q3); range (min–max)	10 (9.0,11.0); (8–16)	10 (9.0, 10.0)	10 (9.0, 9.9)	0.941	
Presence of co-existing cancers (Hong Kong Top 10), n (%)				0.401	
No	557 (97.4)	391 (97.8)	166 (96.5)		
Yes	15 (2.6)	9 (2.3)	6 (3.5)		
Primary tumor					
Tumor location, n (%)				0.852	1 (0.2)
Right	143 (25.0)	101 (25.3)	42 (24.4)		
Left	220 (38.5)	155 (38.8)	65 (37.8)		
Rectum	140 (24.5)	94 (23.6)	46 (26.7)		
Rectum + Colon	69 (12.1)	50 (12.5)	19 (11.0)		
Synchronous colorectal & liver lesion, n (%)				0.345	
No	221 (38.6)	149 (37.3)	72 (41.9)		
Yes	351 (61.4)	251 (62.8)	100 (58.1)		

Table S2 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-453

Table S2 (*continued*)

Variables	Total (n=572)	Training set (n=400)	Testing set (n=172)	P value	Missing data, n (%)
Synchronous colorectal & liver resection, n (%)				0.249	
No	398 (69.6)	272 (68.0)	126 (73.3)		
Yes	174 (30.4)	128 (32.0)	46 (26.7)		
Colorectal CEA, median (Q1, Q3)	9.7 (3.9, 42.0)	10.0 (4.1, 43.5)	8.6 (3.5, 38.9)	0.196	98 (17.1)
ECOG score, n (%)				0.183	255 (44.6)
0	138 (43.5)	93 (43.1)	45 (44.6)		
1	169 (53.3)	120 (55.6)	49 (48.5)		
2	8 (2.5)	5 (2.3)	3 (3.0)		
3	2 (0.6)	0 (0)	2 (2.0)		
Neoadjuvant chemotherapy, n (%)				0.402	8 (1.4)
No	450 (79.8)	311 (78.9)	139 (81.8)		
Yes	114 (20.2)	84 (21.3)	30 (17.6)		
Adjuvant therapy, n (%)				0.334	
No	201 (35.1)	135 (33.8)	66 (38.4)		
Yes	371 (64.9)	265 (66.3)	106 (61.6)		
Surgery admission category, n (%)				0.867	
Emergency	87 (15.2)	62 (15.5)	25 (14.5)		
Elective	485 (84.8)	338 (84.5)	147 (85.5)		
Primary CRC Pathology					
Largest tumor diameter (cm), median (Q1, Q3)	4.0 (3.0, 5.0)	4.0 (3.0, 5.0)	4.0 (3.0, 5.0)	0.479	119 (20.8)
Lymphovascular invasion, n (%)				0.788	92 (16.1)
No	259 (54.0)	181 (54.5)	78 (52.7)		
Yes	221 (46.0)	151 (45.5)	70 (47.3)		
Venous infiltration, n (%)				0.400	285 (49.8)
No	166 (57.8)	109 (55.9)	57 (62.0)		
Yes	121 (42.2)	86 (44.1)	35 (38.0)		
Resection margin, n (%)				1	53 (9.3)
Negative	499 (96.0)	346 (96.1)	153 (95.6)		
Microscopic	20 (3.8)	14 (3.9)	6 (3.8)		
Shortest resection margin, median (Q1, Q3)	3.5 (2.0, 5.0)	3.5 (2.0, 5.4)	3.3 (2.0, 5.0)	0.618	159 (27.8)
Differentiation, n (%)				0.772	43 (7.5)
Well differentiated	18 (3.4)	13 (3.5)	5 (3.2)		
Moderately differentiated	486 (91.9)	338 (91.1)	148 (93.7)		
Poorly differentiated	25 (4.7)	19 (5.1)	6 (3.8)		

Table S2 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-453

Table S2 (*continued*)

Variables	Total (n=572)	Training set (n=400)	Testing set (n=172)	P value	Missing data, n (%)
Number of lymph nodes positive, median (Q1, Q3)	1 (0, 4.0)	2 (0, 4.0)	1 (0, 3.0)	0.163	14 (2.4)
N-stage, n (%)				0.447	14 (2.4)
0	198 (35.5)	134 (34.0)	64 (39.0)		
1–3	218 (39.1)	151 (38.3)	67 (40.9)		
>3	142 (25.4)	105 (26.6)	37 (22.6)		
Colorectal liver metastasis					
Pre-treatment PET scan findings					
Number of tumor nodules, median (Q1, Q3); range (min–max)	1.0 (1.0, 3.0); (0–12)	1.5 (1.0, 3.0)	1.0 (1.0, 3.0)	0.932	78 (13.6)
Tumor location, n (%)				0.613	86 (15.0)
Unilobar	343 (70.6)	235 (69.9)	108 (72.0)		
Bilobar	143 (29.4)	102 (30.4)	41 (27.3)		
Largest tumor diameter (cm), median (Q1, Q3); range (min–max)	2.5 (1.6, 4.0); (0–16.8)	2.4 (1.6, 3.9)	2.6 (1.7, 4.3)	0.368	154 (26.9)
Maximum SUV of lesion, median (Q1, Q3)	6.4 (4.8, 8.9)	6.65 (4.8, 8.9)	6.3 (4.7, 9.1)	0.963	193 (33.7)
Background SUV, median (Q1, Q3)	2.4 (2.0, 2.8)	2.4 (2.0, 2.7)	2.3 (1.9, 2.9)	0.538	433 (75.7)
Extrahepatic metastasis, n (%)				0.687	99 (17.3)
No	423 (89.4)	288 (88.9)	135 (90.6)		
Yes	50 (10.6)	36 (11.1)	14 (9.4)		
Pre-operative investigations					
HBsAg, n (%)				1	260 (45.5)
Negative	286 (91.7)	200 (92.2)	86 (90.5)		
Positive	26 (8.3)	18 (8.3)	8 (8.4)		
HCV Ab, n (%)				1	387 (67.7)
Negative	185 (100.0)	133 (100.0)	52 (100.0)		
Platelet ($x10^9$/L), median (Q1, Q3)	216.0 (175.0, 275.0)	220.0 (172.0, 280.0)	212.0 (175.3, 262.8)	0.383	3 (0.5)
Prothrombin time, median (Q1, Q3)	11.5 (10.8, 12.4)	11.4 (10.7, 12.3)	11.6 (10.8, 12.5)	0.403	21 (3.7)
Albumin (g/L), median (Q1, Q3)	40.0 (36.0, 43.0)	40.0 (35.9, 43.0)	40.0 (36.0, 43.0)	0.761	3 (0.5)
Bilirubin (μmol/L), median (Q1, Q3); range (min–max)	8.2 (6.0, 12.9); (2.0–51.0)	8.5 (6.0, 13.0)	8.0 (6.0, 12.0)	0.518	3 (0.5)
Liver CEA (ng/mL), median (Q1, Q3); range (min–max)	12.0 (4.6, 45.4); (0.7–4,040.0)	11.8 (4.7, 45.2)	13.0 (4.4, 45.0)	0.988	13 (2.3)

Table S2 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-453

Table S2 (*continued*)

Variables	Total (n=572)	Training set (n=400)	Testing set (n=172)	P value	Missing data, n (%)
Liver pathology					
Largest diameter of liver metastasis (cm), median (Q1, Q3); range (min–max)	3.0 (2.0, 4.0); (0.4–18.0)	2.8 (2.0, 4.0)	3.0 (2.0, 4.0)	0.766	7 (1.2)
Number of tumor nodules, median (Q1, Q3); range (min–max)	2 (1.0, 3.0); (0–8)	2 (1.0, 3.0)	2 (1.0, 3.0)		5 (0.9)
Lymphovascular invasion, n (%)				1	412 (72.0)
No	84 (52.5)	57 (51.8)	27 (54.0)		
Yes	76 (47.5)	52 (47.3)	24 (48.0)		
Resection margin, n (%)				0.576	9 (1.6)
Negative	482 (85.6)	335 (85.0)	147 (87.0)		
Microscopic	54 (9.6)	41 (10.4)	13 (7.7)		
Macroscopic	27 (4.8)	18 (4.6)	9 (5.3)		
Shortest resection margin, median (Q1, Q3)	0.5 (0.1, 1.0)	0.5 (0.1, 1.0)	0.5 (0.1, 1.0)	0.914	93 (16.3)
Tumor lobar involvement, n (%)				0.322	2 (0.3)
Unilobar	395 (69.3)	271 (67.9)	124 (72.5)		
Bilobar	175 (30.7)	128 (32.1)	47 (27.5)		
KRAS, n (%)				0.412	147 (25.7)
No mutation	250 (58.8)	168 (56.9)	82 (63.1)		
Mutation	175 (41.2)	125 (42.4)	50 (38.5)		
BRAF, n (%)				1	563 (98.4)
No mutation	8 (88.9)	6 (85.7)	2 (100.0)		
Mutation	1 (11.1)	1 (14.3)	0 (0)		
New Edmondson grading, n (%)				1	456 (79.7)
Well differentiated	2 (1.7)	2 (2.3)	0 (0)		
Moderately differentiated	103 (88.8)	77 (87.5)	26 (92.9)		
Poorly differentiated	9 (7.8)	7 (8.0)	2 (7.1)		
Liver resection					
Laparoscopic, n (%)				1	2 (0.3)
No	455 (79.8)	319 (79.9)	136 (79.5)		
Yes	115 (20.2)	80 (20.1)	35 (20.5)		
Operation, n (%)				0.204	3 (0.5)
Minor	251 (44.1)	167 (42.0)	84 (49.1)		
Major	318 (55.9)	231 (58.0)	87 (50.9)		

Table S2 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-453

Table S2 (*continued*)

Variables	Total (n=572)	Training set (n=400)	Testing set (n=172)	P value	Missing data, n (%)
RFA combined treatment, n (%)				0.505	2 (0.3)
No	518 (90.9)	360 (90.0)	158 (92.9)		
Yes	52 (9.1)	39 (9.8)	13 (7.6)		
Neoadjuvant/Adjuvant therapy					
Neoadjuvant treatment, n (%)				0.871	5 (0.9)
No	378 (66.7)	264 (66.0)	114 (68.3)		
Yes	189 (33.3)	134 (33.5)	55 (32.9)		
Adjuvant chemotherapy, n (%)				0.319	
No	161 (28.1)	118 (29.5)	43 (25.0)		
Yes	411 (71.9)	282 (70.5)	129 (75.0)		
Fong Clinical Risk Score (CRS)					
CEA, n (%)				0.420	
<200	533 (93.2)	370 (92.5)	163 (94.8)		
≥200	39 (6.8)	30 (7.5)	9 (5.2)		
CRC number of lymph nodes positive, n (%)				0.479	
0	212 (37.1)	144 (36.0)	68 (39.5)		
≥1	360 (62.9)	256 (64.0)	104 (60.5)		
Disease free interval (days), n (%)				0.517	
≥365	112 (19.6)	75 (18.8)	37 (21.5)		
<365	460 (80.4)	325 (81.3)	135 (78.5)		
CRLM number of tumor nodules, n (%)				0.841	
≤1	284 (49.7)	197 (49.3)	87 (50.6)		
>1	288 (50.3)	203 (50.8)	85 (49.4)		
CRLM largest tumor diameter, n (%)				0.521	
≤5	492 (86.0)	347 (86.8)	145 (84.3)		
>5	80 (14.0)	53 (13.3)	27 (15.7)		
Fong CRS, median (Q1, Q3)	2 (2.0, 3.0)	2 (1.8, 3.0)	2 (2.0, 3.0)		
Child-Pugh score					
Bilirubin, n (%)				1	3 (0.5)
≤35	567 (99.6)	397 (99.5)	170 (100.0)		
>35	2 (0.4)	2 (0.5)	0 (0)		
Albumin, n (%)				0.524	3 (0.5)
>35	437 (76.8)	303 (75.8)	134 (79.3)		
≤35	132 (23.2)	96 (24.0)	36 (21.3)		

Table S2 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-453

Table S2 (*continued*)

Variables	Total (n=572)	Training set (n=400)	Testing set (n=172)	P value	Missing data, n (%)
Platelet, n (%)				1	3 (0.5)
>100	564 (99.1)	395 (99.0)	169 (99.4)		
≤100	5 (0.9)	4 (1.0)	1 (0.6)		
Nordlinger					
Age group, n (%)				0.331	
≤60	242 (42.5)	175 (43.8)	67 (39.0)		
>60	330 (58.0)	225 (56.3)	105 (61.0)		
Disease-free interval (years), n (%)				0.264	
≥2y	47 (8.3)	29 (7.3)	18 (10.5)		
<2y	525 (92.3)	371 (92.8)	154 (89.5)		
CRLM number of tumor nodules, n (%)				0.080	5 (0.9)
≤3	472 (83.2)	322 (80.9)	150 (88.8)		
>3	95 (16.8)	74 (18.6)	21 (12.4)		
CRLM largest tumor diameter, n (%)				0.523	7 (1.2)
≤5	485 (85.8)	342 (87.0)	143 (83.1)		
>5	80 (14.2)	53 (13.5)	27 (15.7)		
Follow-up					
Post-operative recurrence, n (%)				0.835	
None	163 (28.8)	117 (29.3)	46 (26.7)		
Intrahepatic	70 (12.4)	46 (11.5)	24 (14.0)		
Extrahepatic	89 (15.8)	62 (15.5)	27 (15.7)		
Intrahepatic + extrahepatic	250 (44.2)	175 (43.8)	75 (43.6)		
Follow-up status at conclusion of study, n (%)				0.490	
Alive, disease free	143 (25.3)	99 (24.8)	44 (25.6)		
Alive, with disease (not recurrence)	28 (5.0)	21 (5.3)	7 (4.1)		
Alive, with recurrence	60 (10.6)	48 (12.0)	12 (7.0)		
Alive, details unavailable	5 (0.9)	3 (0.8)	2 (1.2)		
Dead	327 (57.9)	223 (55.8)	104 (60.5)		
Default	9 (1.6)	6 (1.5)	3 (1.7)		
Hospital mortality, n (%)				0.570	1 (0.2)
No	556 (97.4)	388 (97.0)	168 (98.2)		
Yes	15 (2.6)	12 (3.0)	3 (1.8)		

Table S2 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-453

Table S2 (*continued*)

Variables	Total (n=572)	Training set (n=400)	Testing set (n=172)	P value	Missing data, n (%)
Cause of death, n (%)				1	245 (42.8)
Dead from other causes/disease	71 (21.7)	49 (22.0)	22 (21.2)		
Dead from original disease (CRLM) or recurrence or metastasis	256 (78.3)	175 (78.5)	81 (77.9)		
Status, n (%)				0.304	
Censored (alive at the end of the study or was lost to follow up)	243 (42.5)	176 (44.0)	67 (39.0)		
Dead	329 (57.5)	224 (56.0)	105 (61.0)		
Recurrence-free survival (RFS), median (Q1, Q3)	419.0 (186.5, 1,159.8)	437.5 (179.8, 1,182.3)	386.5 (201.0, 1,085.8)	0.807	
Overall survival (OS), median (Q1, Q3)	1,169.0 (771.5, 1,893.0)	1,169.0 (748.8, 1,922.3)	1,167.0 (824.8, 1,844.3)	0.999	
Disease free interval (DFI), median (Q1, Q3)	0 (0, 218.0)	0 (0, 204.8)	0 (0, 289.0)	0.182	

QMH, Queen Mary Hospital; QEH, Queen Elizabeth Hospital; TMH, Tuen Mun Hospital; PWH, Prince of Wales Hospital; HBsAg, Hepatitis B surface antigen; HCV Ab, Hepatitis C antibody; CEA, carcinoembryonic antigen; ECOG, Eastern Cooperative Oncology Group (ECOG) performance status; PET, positron emission tomography; SUV, standardizsed uptake value; RFA, radiofrequency ablation; DFI, disease-free interval (time from primary tumor diagnosis to colorectal liver metastasis).

https://dx.doi.org/10.21037/hbsn-21-453

Table S3 Cox proportional hazards univariate analysis for OS and RFS

Variable	OS			RFS		
	HR	Estimate (CI)	P value	HR	Estimate (CI)	P value
CRLM number of tumor nodules	3.761	0.127 (0.061–0.193)	<0.001	5.470	0.182 (0.117–0.247)	<0.001
Fong CRS	3.888	0.206 (0.102–0.310)	<0.001	5.216	0.275 (0.172–0.378)	<0.001
Nordlinger score - CRLM number of tumor nodules	–4.123	–0.480 (–0.709, –0.252)	<0.001	4.841	0.661 (0.393–0.928)	<0.001
CRLM adjuvant chemotherapy	3.397	0.080 (0.034–0.125)	<0.001	–4.352	–0.507 (–0.735, –0.279)	<0.001
Fong CRS - CRLM number of tumor nodules	3.262	0.443 (0.177–0.710)	0.001	3.860	0.429 (0.211–0.647)	<0.001
CRLM largest diameter of liver metastasis	3.334	0.580 (0.239–0.921)	0.002	3.580	0.083 (0.038–0.129)	<0.001
CRC N-stage	3.169	2.271 (0.866–3.676)	0.002	3.268	0.233 (0.093–0.372)	0.001
PET - number of tumor nodules	3.135	0.224 (0.084–0.365)	0.002	3.356	0.131 (0.054–0.207)	0.002
PET - extrahepatic metastasis	2.946	0.126 (0.042–0.210)	0.003	3.120	0.524 (0.195–0.853)	0.003
CRLM neoadjuvant treatment	2.802	0.031 (0.009–0.052)	0.005	2.909	0.335 (0.109–0.560)	0.004
PET - tumor location	2.692	0.313 (0.085–0.540)	0.008	2.922	0.388 (0.128–0.649)	0.005
Fong CRS - CEA group	2.466	0.273 (0.056–0.491)	0.014	2.807	0.556 (0.168–0.944)	0.005
CRC lymphovascular invasion	2.394	0.611 (0.111–1.112)	0.018	2.789	0.332 (0.099–0.566)	0.006
CRLM tumor lobar involvement	–2.339	–0.355 (–0.652, –0.057)	0.020	2.711	0.317 (0.088–0.545)	0.007
CRC number of lymph nodes positive	2.246	0.086 (0.011–0.160)	0.028	2.590	0.027 (0.007–0.048)	0.010
CRLM liver resection - laparoscopic	2.165	0.315 (0.030–0.600)	0.042	–2.507	–0.380 (–0.677, –0.083)	0.013
CRC resection margin	2.015	0.236 (0.006–0.465)	0.045	2.479	0.611 (0.128–1.095)	0.014
Fong CRS - DFI	1.995	0.394 (0.007–0.782)	0.047	2.363	0.347 (0.059–0.635)	0.019
CRLM RFA combined treatment	2.395	0.577 (0.105–1.048)	0.054	2.330	0.418 (0.066–0.769)	0.020
Fong CRS - CRC number of lymph nodes positive	1.955	0.262 (–0.001–0.524)	0.055	2.279	0.267 (0.037–0.496)	0.023
PET - largest tumor diameter	1.857	0.272 (–0.015–0.560)	0.064	2.286	0.052 (0.007–0.097)	0.024
Charlson Comorbidity Score	1.809	0.211 (–0.018–0.440)	0.071	2.115	0.091 (0.007–0.175)	0.035
Fong CRS - CRLM number of tumor nodules	1.748	0.009 (–0.001–0.020)	0.081	2.025	0.304 (0.010–0.598)	0.044
CRLM lymphovascular invasion	1.744	0.031 (–0.004–0.065)	0.086	2.115	0.417 (0.031–0.803)	0.067
Bilirubin (>35 μmol/L) pre-hepatectomyn	1.677	0.192 (–0.032–0.417)	0.095	1.634	1.161 (–0.231–2.553)	0.103
Colorectal CEA	1.659	0.249 (–0.045–0.543)	0.098	1.600	0.000 (0.000–0.000)	0.111
CRC synchronous lesion	1.733	0.273 (–0.036–0.581)	0.104	1.493	0.170 (–0.053–0.393)	0.137
KRAS	1.623	0.060 (–0.013–0.133)	0.106	1.560	0.237 (–0.061–0.535)	0.138
Platelet	1.549	0.169 (–0.045–0.383)	0.122	–1.358	–0.001 (–0.002–0.000)	0.176
CRC neoadjuvant chemotherapy	1.484	0.165 (–0.053–0.384)	0.139	1.305	0.180 (–0.090–0.450)	0.193
Age	–1.450	–0.083 (–0.195–0.029)	0.148	1.277	0.007 (–0.004–0.017)	0.202

Table S3 (*continued*)

https://dx.doi.org/10.21037/hbsn-21-453

Table S3 (*continued*)

Variable	Overall survival (OS)			Recurrence-free survival (RFS)		
	HR	Estimate (CI)	P value	HR	Estimate (CI)	P value
Hospital	−1.409	−0.094 (−0.225–0.037)	0.161	−1.273	−0.059 (−0.150–0.032)	0.204
Liver resection - operation	1.392	0.248 (−0.101–0.598)	0.165	1.267	0.141 (−0.077–0.359)	0.206
Drinking	1.262	0.012 (−0.006–0.030)	0.208	1.212	0.114 (−0.070–0.299)	0.227
Smoking	1.251	0.140 (−0.080–0.361)	0.212	1.198	0.104 (−0.066–0.275)	0.234
Sex	−1.117	−0.130 (−0.358–0.098)	0.265	−1.154	−0.134 (−0.362–0.094)	0.249
CRLM resection margin	1.001	0.023 (−0.022–0.067)	0.318	1.123	0.120 (−0.090–0.330)	0.263
PET background SUV	0.938	0.000 (0.000–0.000)	0.349	−1.222	−0.217 (−0.565–0.131)	0.269
Nordlinger score - DFI	0.919	0.081 (−0.092–0.255)	0.361	1.097	0.228 (−0.180–0.636)	0.274
Prothrombin time	0.815	0.000 (0.000–0.000)	0.416	1.019	0.037 (−0.034–0.109)	0.309
CRLM shortest resection margin	0.807	0.168 (−0.240–0.576)	0.420	−1.016	−0.069 (−0.202–0.064)	0.311
DFI	−0.873	−0.192 (−0.621–0.239)	0.425	−0.981	0.000 (0.000–0.000)	0.327
CRC tumor location	0.824	0.021 (−0.029–0.071)	0.433	−0.956	−0.055 (−0.168–0.058)	0.340
CRC venous invasion	−0.773	−0.001 (−0.002–0.001)	0.440	0.948	0.136 (−0.145–0.416)	0.353
PET maximum SUV of lesion	0.724	0.147 (−0.250–0.543)	0.470	0.922	0.016 (−0.018–0.050)	0.359
Liver CEA	0.688	0.090 (−0.166–0.346)	0.492	0.903	0.000 (0.000–0.000)	0.367
ECOG score	−0.638	−0.006 (−0.026–0.013)	0.524	0.901	0.092 (−0.108–0.293)	0.370
CRC shortest resection margin	0.473	0.114 (−0.358–0.585)	0.644	0.889	0.020 (−0.024–0.063)	0.392
Bilirubin	0.409	0.048 (−0.181–0.277)	0.683	0.820	0.007 (−0.010–0.025)	0.413
Synchronous colorectal & liver resection	−0.389	0.000 (0.000–0.000)	0.698	0.684	0.082 (−0.152–0.315)	0.495
Fong CRS - age group	0.353	0.016 (−0.074–0.107)	0.725	0.531	0.060 (−0.161–0.280)	0.596
Child Pugh Score - albumin	0.273	0.043 (−0.268–0.354)	0.791	0.500	0.065 (−0.191–0.321)	0.618
CRC adjuvant therapy	0.263	0.024 (−0.156–0.205)	0.793	0.447	0.052 (−0.177–0.281)	0.655
CRC New Edmondson Grade	0.261	0.088 (−0.575–0.751)	0.794	−0.391	−0.080 (−0.482–0.322)	0.696
Albumin	0.257	0.092 (−0.609–0.792)	0.802	−0.367	−0.004 (−0.024–0.017)	0.714
HBsAg	0.191	0.022 (−0.201–0.245)	0.849	0.259	0.059 (−0.384–0.501)	0.799
CRC largest tumor diameter	0.170	0.023 (−0.246–0.293)	0.865	−0.198	−0.006 (−0.067–0.055)	0.844
CRC surgery admission category	0.127	0.074 (−1.064–1.212)	0.899	−0.106	−0.016 (−0.320–0.287)	0.916
CRLM New Edmondson Grade	0.120	0.004 (−0.057–0.065)	0.905	0.078	0.013 (−0.320–0.347)	0.940
Child Pugh - platelet	−0.055	−0.009 (−0.312–0.295)	0.956	0.074	0.043 (−1.095–1.181)	0.941
Presence of co-existing cancers (Hong Kong Top 10)	−0.050	−0.005 (−0.213–0.202)	0.960	0.057	0.019 (−0.644–0.682)	0.955
BRAF	−0.005	−0.001 (−0.234–0.233)	0.996	0.024	0.010 (−0.818–0.838)	0.981
HCV Ab	NA due to the large amount of missing data					

CRLM, colorectal liver metastasis; CRS, clinical risk score; CRC, colorectal cancer; HBsAg, Hepatitis B surface antigen; HCV Ab, Hepatitis C antibody; CEA, carcinoembryonic antigen; ECOG, Eastern Cooperative Oncology Group (ECOG) performance status; PET, positron emission tomography; SUV, standardized uptake value; RFA, radiofrequency ablation; DFI, disease-free interval (time from primary tumor diagnosis to colorectal liver metastasis).

https://dx.doi.org/10.21037/hbsn-21-453