OXFORD

BJS Open, 2021, zrab104
DOI: 10.1093/bjsopen/zrab104
Original Article

Composite Score: prognostic tool to predict survival in patients undergoing surgery for colorectal liver metastases

Petter Frühling[1,*], Jozef Urdzik[1], Cecilia Strömberg[2] and Bengt Isaksson[1]

[1]Department of Surgical Sciences, Uppsala University, Uppsala, Sweden
[2]Division of Surgery, Department for Clinical Science, Intervention and Technology (CLINTEC), Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden

*Correspondence to: Dr P. Frühling, Department of Surgical Sciences, Uppsala University, Akademiska sjukhuset ing 70 1 tr, 751 85 Uppsala, Sweden (e-mail: petter.fruhling@surgsci.uu.se)

Abstract

Background: Several existing scoring systems predict survival of patients with colorectal liver metastases. Many lack validation, rely on old clinical data, and have been found to be less accurate since the introduction of chemotherapy. This study aimed to construct and validate a clinically relevant preoperative prognostic model for patients with colorectal liver metastases.

Methods: A predictive model with data available before surgery was developed. Survival was analysed by Cox regression analysis, and the quality of the model was assessed using discrimination and calibration. The model was validated using multifold cross-validation.

Results: The model included 1212 consecutive patients who underwent liver resection for colorectal liver metastases between 2005 and 2015. Prognostic factors for survival included advanced age, raised C-reactive protein level, hypoalbuminaemia, extended liver resection, larger number of metastases, and midgut origin of the primary tumour. A Composite Score was developed based on the prognostic variables. Patients were classified into those at low, medium, and high risk. Survival differences between the groups were significant; median overall survival was 87.4 months in the low-risk group, 50.1 months in the medium-risk group, and 22.6 months in the high-risk group. The discriminative performance, assessed by the concordance index, was 0.71, 0.67, and 0.67 respectively at 1, 3, and 5 years. Calibration, assessed graphically, was close to perfect. A multifold cross-validation of the model confirmed its internal validity (C-index 0.63 *versus* 0.62).

Conclusion: The Composite Score categorizes patients into risk strata, and may help identify patients who have a poor prognosis, for whom surgery is questionable.

Introduction

Liver resection is the most effective treatment option for patients with colorectal liver metastases (CRLMs). Unfortunately, only a minority of patients have resectable metastases at the time of diagnosis. The prognosis for patients with untreated CRLMs is poor. Synchronous CRLMs often signify worse cancer biology and shorter expected overall survival. The prognosis for patients with CRLMs has improved over recent decades. In a systematic review and meta-analysis[1] from 2012, which included 35 studies, the 5-year overall survival rate was approximately 40 per cent.

Several prognostic scoring systems exist in the field of CRLMs, many of which use two to seven different clinical parameters to calculate a risk score[2–7]. Limitations of these scoring systems are manifold. First, as they were developed before the introduction of neoadjuvant chemotherapy, their role as reliable predictive tools is questionable[8,9]. Second, several of these scoring systems are cumbersome in their work-up, and rely on the pathology report for information on lymph node status[3] and serosal invasion of the primary tumour[6], which precludes their use in the preoperative setting. Finally, few of the scoring systems have been validated adequately.

A great deal of research has gone into identifying patient- and tumour-specific variables associated with survival in patients with CRLMs. Most studies agree on the importance of tumour size (for example, 5 cm or less, or 5 cm and over) and number of metastases (single or multiple)[10,11]. In a randomized study by Nordlinger and colleagues[12,13], no overall survival difference was seen in the chemotherapy group compared with surgery alone.

In the past two decades, several studies have shed light on the interaction between inflammation and cancer. Epidemiological data support the theory that systemic inflammation, as measured by a raised C-reactive protein (CRP) level and hypoalbuminemia, is associated with worse outcome in advanced cancers[14–18]. A score based on CRP and albumin was developed by Forrest and co-workers[19,20], called the Glasgow Prognostic Score. In recent years, several studies have examined the role of molecular markers[21–23], both as prognostic and predictive factors, as well as

Received: May 23, 2021. **Accepted:** September 16, 2021

their role in assessing chemosensitivity and toxicity[24–27]. The aim of the present study was to develop, and validate, a simple predictive model for patients with resectable CRLMs, based on preoperative variables.

Methods

Study population

This study used data from all consecutive patients who underwent surgery with curative intent for CRLMs between 1 January 2005 and 31 December 2015 at Karolinska and Uppsala University Hospitals. All patients had CRLMs verified radiologically or by biopsy. Clinicopathological factors were collected retrospectively from prospectively recorded databases. Data retrieved included: age, sex, primary tumour site, embryonic origin of primary, BMI, ASA physical status classification, timing of metastasis (synchronous/metachronous), and preoperative albumin and CRP values. All patients were discussed in a liver-specific multidisciplinary meeting, with dedicated radiologists, oncologists, and liver surgeons present, and were considered candidates for liver surgery with curative intent.

Patients treated with neoadjuvant chemotherapy were considered candidates for surgery if radiological evaluation after four or six cycles of chemotherapy showed an objective response or stable disease. To keep the composite model simple, molecular markers and chemotherapy treatment were not included in the construction of the model. The study was approved by the regional ethical review board.

Prognostic factors

Prognostic factors for patients with CRLMs in the literature were considered. They included demographic information (age and sex), tumour characteristics of the liver metastasis (tumour size and number of metastases), primary location of tumour (colon, rectum), embryonic origin, and systemic inflammatory markers, such as serum albumin and CRP[3,4,28–30].

Follow-up after surgery

All patients were followed up in the outpatient clinic 4–6 weeks after surgery. Thereafter, patients were followed up according to the national guidelines for colorectal cancer with liver metastases, which as a minimum entailed yearly contrast-enhanced CT of abdomen and thorax. Patients were followed up routinely for at least 5 years. In the present study, survival and follow-up was calculated from the date of surgery for CRLMs. End of follow-up was 13 July 2020; this was the date on which survival status was obtained from the Total Population Register, which is updated continuously and linked to each patient's personal identity number in the electronic medical notes.

Statistical analysis

Categorical variables are presented as numbers with percentages, and continuous data as median (i.q.r.). Categorical variables were compared using the χ^2 test, and the Kruskal–Wallis test was used for continuous data. Overall survival was defined as the interval between the date of surgery and death from any cause. Time was censored at the last follow-up for patients who were still alive. Survival curves were constructed using the Kaplan–Meier method, and differences in survival assessed with the log rank test. The median follow-up time was calculated using the reversed Kaplan–Meier method[31]. A Cox proportional hazards model was used to assess the effect of the prognostic factors. Prognostic factors were included in the multivariable model based on clinical and statistical significance. As data were missing for some variables, multiple imputation was used to impute missing values, according to established practice. The method of imputation was set at the automatic setting in SPSS, which uses a combination of the monotonic and Markov chain Monte Carlo methods. To reduce sampling variability in the imputation process, a total of 25 imputed data sets were generated[32–34].

In the Cox model, backward elimination was used to assess the relationship between relevant clinicopathological variables and overall survival. Variables with $P < 0.050$ in the univariable analysis were included in the multivariable analysis. The Composite Score was created based on the multivariable analysis, and included the β-coefficients of the significant variables. Each patient was thereafter assigned an accumulated β-coefficient score (*Fig. S1*). Survival estimates were reported as hazard ratios (HRs) with 95 per cent confidence intervals. The performance of the model was assessed in terms of discrimination and calibration. Discrimination reflects the model's ability to discriminate patients who have the outcome (in this case death) from those without the outcome (alive). Calibration refers to the agreement between the observed outcome (estimated overall survival) and predicted overall survival[35]. The discrimination of the Composite Score was assessed by receiver operating characteristic curve (ROC) analysis, with overall survival as outcome. For a

Table 1 Baseline characteristics

	No. of patients (n = 1212)*
Age (years)†	67 (59–73)
> 70	360 (29.7)
> 80	57 (4.7)
Sex	
Men	748 (61.7)
Women	464 (38.3)
BMI (kg/m^2)†	26 (23–28)
ASA fitness grade	
I	109 (9.0)
II	739 (61.0)
III	352 (29.0)
IV	12 (1.0)
Primary tumour location	
Colon	705 (58.2)
Rectum	498 (41.1)
Both	9 (0.7)
Embryonic origin	
Midgut	254 (21.0)
Hindgut	947 (78.1)
Unclear	11 (0.9)
Timing of metastasis	
Synchronous	700 (57.8)
Metachronous	512 (42.2)
Tumour size (mm)†	25 (20–40)
No. of tumours	
1–2	477 (39.4)
3–5	589 (48.6)
> 5	146 (12.0)
Resection type (no. of liver segments)	
Minor (< 3)	670 (55.3)
Major (3–4)	413 (34.1)
Extended (> 4)	129 (10.6)
Laboratory data	
C-reactive protein (mg/l)	
<10	876 (72.3)
10 to < 50	301 (24.8)
≥ 50	35 (2.9)
Albumin < 35 g/l	410 (33.8)

*With percentages in parentheses unless indicated otherwise; †values are median (i.q.r.).

Table 2 Univariable and multivariable analyses of prognostic scores after imputation, and β-coefficient for each variable

	Univariable analysis		Multivariable analysis		β-coefficient
	Hazard ratio	P	Hazard ratio	P	
Men	1.08 (0.94, 1.25)	0.285			
Age (years)					
< 70	1.00 (reference)		1.00 (reference)		
≥ 70–80 years	1.29 (1.11, 1.51)	0.001	1.33 (1.13, 1.57)	< 0.001	0.30
> 80 years	2.58 (1.94, 3.42)	< 0.001	3.28 (2.47, 4.39)	< 0.001	1.16
Midgut location	1.22 (1.18, 1.26)	0.019	1.20 (1.10, 1.54)	0.039	0.18
Synchronous	1.16 (1.01, 1.34)	0.039	1.06 (0.87, 1.30)	0.547	
C-reactive protein (mg/l)					
< 10	1.00 (reference)		1.00 (reference)		
10– <50	1.48 (1.15, 1.89)	0.002	1.17 (1.01, 1.36)	0.010	0.26
≥ 50	1.66 (1.04, 2.66)	0.033	1.75 (1.14, 2.69)	0.033	0.56
Albumin (g/l)					
≥ 35	1.00 (reference)		1.00 (reference)		
< 35	1.33 (1.13, 1.58)	0.001	1.30 (1.04, 1.63)	0.022	0.14
Perioperative ablation	1.81 (1.29, 2.54)	0.001	1.12 (0.69, 1.83)	0.651	
Two-stage resection	1.71 (1.06, 2.77)	0.028	0.91(0.51, 1.63)	0.757	
Extended resection	1.60 (1.33, 1.92)	< 0.001	1.54 (1.16, 2.05)	0.003	0.26
Portal vein embolization	1.19 (0.86, 1.66)	0.303			
Liver metastasis > 50 mm	1.23 (1.03, 1.47)	0.020	1.06 (0.83, 1.35)	0.662	
No. of metastases					
1	1.00 (reference)		1.00 (reference)		
2–5	1.27 (1.10, 1.49)	0.002	1.35 (1.15, 1.59)	< 0.001	0.30
> 5	1.99 (1.59, 2.48)	< 0.001	2.17 (1.71, 2.74)	< 0.001	0.76

Values in parentheses are 95 per cent confidence intervals.

binary outcome, ROC analysis yields an identical result to concordance statistic (C-statistic), which will be reported in the study[36].

Model fit was assessed by visual comparison of Kaplan–Meier plots. Model calibration was assessed with calibration plots, which depicted predicted *versus* observed 3-year overall survival. Perfect predictions should lie along the 45° line. Validation of the C-statistic of the model was performed using a multifold cross-validation, which is an extension of split-sample validation. The data set was split randomly into five *k*-subsets, each containing 20 per cent of the patients. The risk score was thereafter calculated for one of the subsets, and validated on the remaining subsets ($k-1$). This procedure was repeated randomly for each of the *k*-subsets. An average of all risk score estimates and measures of model performance (C-statistic) was thereafter calculated. All analyses were carried out with SPSS® version 27 (IBM, Armonk, NY, USA) and R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria).

Supplementary analyses

The performance of the Composite Score was assessed both on the original data set (before imputation), and after imputation. Visual illustrations of the risk score (*Fig. S1*), and the probability to survive 1-, 3-, and 5-years (*Fig. S2*) are provided in the supplementary material. Moreover, further information about missing data (*Table S1*), Cox regression analyses, and discrimination and calibration before and after imputation can be found in the supplementary material (*Figs S3-S4*, *Tables S3-S4*, and *Figs S5-13*).

Results

During the study interval, a total of 1212 consecutive patients underwent liver resection. Baseline characteristics of the patients are shown in *Table 1*. The median age at surgery was 67 years, a majority had an ASA fitness grade of II or III (1091, 90.0 per cent), and most were men (748, 61.7 per cent). The location of the primary tumour was the colon in 58 per cent of patients, and the rectum in 41 per cent. Nine patients had a synchronous tumour in both the rectum and colon. The median follow-up time was 108 (i.q.r. 103–112) months. At the end of follow-up, 769 patients had died (63.4 per cent), and 443 were censored. The fact that relatively few patients were censored adds stability to the model. *Table S1* shows the number and percentages of variables that were missing. Of the six variables included in the model, four had less than 1 per cent missing values. Values for preoperative CRP and albumin were missing for more than 20 per cent of patients, and it seems plausible to assume that these were missing at random.

Univariable and multivariable analyses

In univariable analyses, synchronous CRLMs, perioperative ablation, two-stage resection, and liver metastasis larger than 50 mm were identified as variables associated with worse overall survival (*Table 2*). In the multivariable analyses, however, these variables were not significant. The multivariable analyses identified six factors associated with worse overall survival, including advanced age (over 70 to 80 years, and more than 80 years), albumin level below 35 g/l, raised CRP level, midgut origin of primary, extended resection (more than 4 liver segments), and larger number of metastases. Of these, especially age over 80 years (HR 3.28, 95 per cent c.i. 2.47 to 4.39), CRP level over 50 mg/l (HR 1.75, 1.14 to 2.69), and more than five metastases (HR 2.17, 1.71 to 2.74) were strongly associated with worse overall survival (*Table 2*).

Stratification of patients according to Composite Score

For each patient, a preoperative Composite Score was calculated by summing their β-coefficients generated in the multivariable analyses. The β-coefficient for each variable is shown in *Table 2*. Baseline characteristics for the three risk groups are shown in *Table S2*. The Composite Score was calculated for all patients

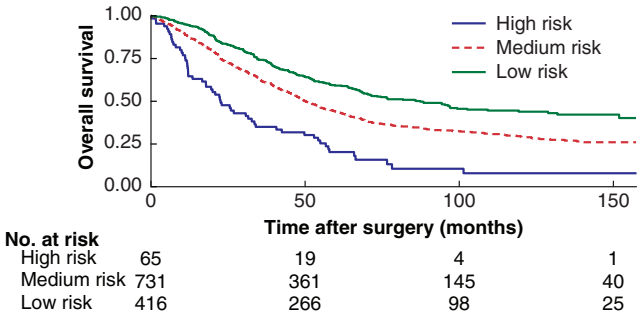

Fig. 1 Overall survival according to risk group

Median overall survival was 87.4 (95 per cent c.i. 69.0 to 119.3) months in the low-risk group, 50.1 (46.8 to 55.2) months in the medium-risk group, and 22.6 (17.8 to 34.0) months in the high-risk group. $P < 0.001$ (log rank test).

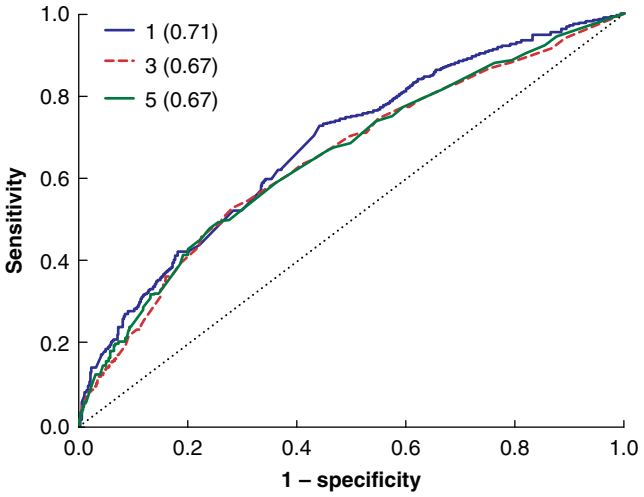

Fig. 2 Receiver operating characteristic (ROC) curves for the model at 1, 3, and 5 years

Area under the curve 0.71, 0.67, and 0.67 at 1, 3, and 5 years respectively.

included in the cohort, as missing values were imputed according to established practice (*Tables S3* and *S4*). A visual illustration of the score is provided in *Fig. S1*. For clinical use, patients were classified into three risk groups: those at low, medium, and high risk. These risk groups were created based on each patient's probability of surviving for 3 years. Patients with an estimated probability of surviving of below 40 per cent were classified as high risk; those with an estimated probability of surviving of between 41 and 70 per cent were classified as medium risk; and patients with an estimated probability of surviving of more than group 70 per cent were classified as low risk. The cut-off β-coefficient values for the three risk groups were: over 1.24 (high risk), 0.27–1.24 (medium risk), and below 0.27 (low risk). The cut-off values were based on the distribution of patients surviving for 3 years (*Fig. S2*), and the aim was to capture patients with both a very short expected survival probability, and those with a considerably better prognosis. More information on the distribution of patients according to the Composite Score, and their probability of surviving at least 3 years, can be obtained from *Figs. S1* and *S2*. A Kaplan–Meier curve was constructed based on these three risk groups (*Fig. 1*). Median overall survival according to the Composite Score was 87.4 months for the low-risk group, 50.1 months for the medium-risk group, and 22.6 months for the high-risk group.

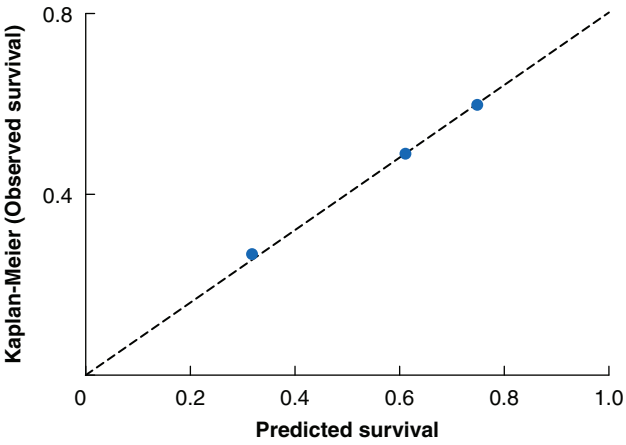

Fig. 3 Calibration plot showing observed survival compared with survival predicted by composite model

Symbols represent the probability of surviving for 3 years according to the three risk groups (low, medium, and high). Perfect calibration follows the diagonal line.

Performance of Composite Score

The composite model showed slightly better discrimination at 1 year than at 3 and 5 years (*Fig. 2*). The Composite Score outperformed the Tumour Burden Score (TBS) in discrimination (C-statistic 0.67 *versus* 0.58). A comparison of the discrimination between the present model and the TBS is provided in *Figs. S6* and *S7*. Visual inspection of the Kaplan–Meier curves (*Fig. 1*) confirmed wide separation between them, which suggests good discrimination[37]. Because the Composite Score did not include genetic information or details regarding the primary tumour, such as primary lymph node status, a comparison could not be made with scoring systems that include these variables. The calibration of the model was assessed at 3 years, and was near to perfect (*Fig. 3*). *Figures S3* and *S4* show other ways of depicting calibration, based on either equal sized groups (*S4*). A multifold cross-validation of the model confirmed the model's internal validity (C-index 0.63 *versus* 0.62).

Comparison of observed and imputed data

A comparison of the observed (original) and imputed data set is provided in *Figs. S5–S6*, and *S8–S13*; the results were consistent with those of the main analyses.

Discussion

In the management of cancer, an accurate predictive model can be used in the provision of effective surveillance. Moreover, it can guide the development of adjuvant strategies and in the planning of future clinical trials, as well as being used in counselling of patients[38–41]. In the era of personalized medicine, the need to develop more accurate predictive tools is paramount.

In this large retrospective study, a predictive model was constructed for patients who underwent liver resection for CRLMs. Liver resection is the only treatment modality that can offer long-term survival for patients with CRLMs[42,43]. The number of patients with CRLM that is considered resectable is increasing. This puts great pressure on clinicians to develop evidence-based selection criteria that allow better discrimination of patients who would benefit most from surgery. To meet this challenge, several studies have been published with the aim of constructing prognostic scores based on tumour morphological factors and

patient-dependent factors. Unfortunately, many of these prognostic scores rely on old patient data, and often lack adequate validation. Moreover, given that treatment in the field of surgical oncology is constantly evolving, their reliability can be questioned[44,45]. Interestingly, Sasaki and colleagues[46] recently published a study concerning a new prognostic tool, the TBS. This model uses a combination of tumour size and number of liver metastases to predict overall survival. The discrimination of the present model was, according to the C-statistic, 0.71 at 1 year, and 0.67 at 3 and 5 years. This is similar to (or marginally better than) values reported by Sasaki *et al.*[46], but the present model outperformed the TBS (C-statistic 0.58) in the present study.

In the past few years, three different models for patients with CRLMs have been developed that include *RAS* mutational status[21], *KRAS* mutational status[22], and alterations in the RAS–RAF pathway or SMAD family[23]. In the largest of these studies by Brudvik and colleagues[21], a modified clinical score for patients with CRLMs was developed, which initially relied on node-positive primary, disease-free survival, more than one liver metastasis, largest liver tumour over 50 mm, carcinoembryonic antigen (CEA) level, and *RAS* mutation. Only patients with data on all six variables were included in the model. Commendably, the authors decided to exclude three of these variables because they did not add any prognostic value to the model, and ended up with only three variables: primary tumour lymph node status, size of largest CRLM, and *RAS* mutation (based on resected specimen). A limitation of this study was the lack of information on the excluded patients, which raises a question of selection bias. In addition, calibration and discrimination was not reported for the whole model, and there was limited information with regard to how validation was carried out.

Patients with data on CEA, primary tumour lymph node status, TBS, extrahepatic disease, and KRAS mutational status were included in the retrospective study by Margonis *et al.*[22]. The authors reported discrimination (C-index 0.645), but did not provide information on the excluded patients, calibration, or how the external validation was performed. The short follow-up of 30.5 months meant that many patients were censored, which potentially adds instability to the model. The model provides three risk groups (low, moderate, and high), but the Kaplan–Meier plot gives the visual impression that the curves for low and moderate risk cross each other, suggesting that there is no difference in survival between these groups. Importantly, in a recent editorial, Margonis *et al.*[47] reflected on the complexity of tumour biology, and the fact that there is a risk of overestimating the importance of single gene-based biomarkers.

In a smaller study by Lang and colleagues[23], cancer-related genes were analysed retrospectively in 139 patients by next-generation sequencing. This study showed that alterations in the RAS–RAF pathway, and in the SMAD family have prognostic significance. Limitations of this study included the short follow-up of 34.5 months, and the lack of information regarding the selection of patients for inclusion in the study.

The ambition of these three retrospective studies to include molecular profiling is commendable. However, what they all illustrated is that routine advanced gene profiling is still not part of the regular work-up for most patients with CRLMs. As extended molecular profiling becomes an integral part of daily clinical routine, it may help in risk stratification in the future. Moreover, these studies rely on data that are often not available before operation, such as lymph node status.

Importantly, the Composite Score is the first clinical score to incorporate CRP and albumin, in conjunction with stratified advanced age, extended liver resection, and tumour characteristics (number of tumours and embryonic origin of primary). Oncological treatment is not included in the Composite Score, because there is wide variation in the indications and timing between centres, and in the choice of oncological agents. In addition, it is difficult to include the impact of number of cycles administered in a meaningful way, and to include tumour response in a straightforward way[48].

Previous risk scores have generally been cumbersome, and relied on information regarding primary tumour lymph node status, had short follow-up, and required advanced genetic profiling which is still not common practice in routine settings. Unlike many other published scoring systems, the aim of the Composite Score was to detect individuals who would benefit the least from surgery. The model was able to identify a group of patients (high risk) with a median overall survival of 22.6 months, which is worse than reported for patients with CRLMs receiving palliative care[49,50]. This highlights the importance of recognizing that a patient with radiologically resectable liver metastases may perhaps not benefit from surgery. By combining number of tumours, preoperative albumin and CRP levels, advanced age, embryonic origin of the primary, and extended liver resection, a new prognostic tool was constructed. To maintain the simplicity of the model for the clinician to use, molecular markers were excluded, as well as the role of chemotherapy. The long follow-up resulted in there being few censored patients, which adds stability to the analysis.

Three distinct risk groups were created in the Composite Score. To identify patients with more extreme prognoses and to group together those with similar prognoses, the risk groups were of unequal size. The aim was to minimize the loss of information that occurs in grouping. The Composite Score showed a higher degree of discrimination in its estimate of survival after 1 year compared to 3 and 5 years. This is likely to be explained by changes in conditional survival over time, and competing risks that affect patients who live for a long time after initially being diagnosed with CRLMs. The calibration at 3 years was near to perfect, and the multifold cross-validation confirmed the model's internal validity. Importantly, for a new prediction model to be generalizable and introduced into clinical practice, it needs to be validated externally. This requires that the model is evaluated in new patients in a different population[51,52].

Limitations of the present study include the fact that there were missing values for important variables. If values are not missing at random, there is a risk of introducing bias. To address this, multiple imputation was used. A complete-case analysis was performed, which compared the original data set (with complete information regarding all variables) and the imputed data sets. The results of these analyses were consistent. If the Composite Score can be validated further, the aim is to create an online risk calculator that may help clinicians in the selection of surgical candidates.

Acknowledgements

The authors acknowledge J. Bring at Statisticon, Statistics and Research for assistance and guidance with statistical analyses.

Disclosure. The authors declare no conflict of interest.

Supplementary material

Supplementary material is available at *BJS Open* online.

References

1. Jones RP, Jackson R, Dunne DF, Malik HZ, Fenwick SW, Poston GJ *et al.* Systematic review and meta-analysis of follow-up after hepatectomy for colorectal liver metastases. *Br J Surg* 2012;**99**:477–486.
2. Kattan MW, Gonen M, Jarnagin WR, DeMatteo R, D'Angelica M, Weiser M *et al.* A nomogram for predicting disease-specific survival after hepatic resection for metastatic colorectal cancer. *Ann Surg* 2008;**247**:282–287.
3. Fong Y, Fortner J, Sun RL, Brennan MF, Blumgart LH. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. *Ann Surg* 1999;**230**:309–318.
4. Nordlinger B, Guiguet M, Vaillant JC, Balladur P, Boudjema K, Bachellier P *et al.* Surgical resection of colorectal carcinoma metastases to the liver. A prognostic scoring system to improve case selection, based on 1568 patients. Association Francaise de Chirurgie. *Cancer* 1996;**77**:1254–1262.
5. Iwatsuki S, Dvorchik I, Madariaga JR, Marsh JW, Dodson F, Bonham AC *et al.* Hepatic resection for metastatic colorectal adenocarcinoma: a proposal of a prognostic scoring system. *J Am Coll Surg* 1999;**189**:291–299.
6. Nagashima I, Takada T, Matsuda K, Adachi M, Nagawa H, Muto T *et al.* A new scoring system to classify patients with colorectal liver metastases: proposal of criteria to select candidates for hepatic resection. *J Hepatobiliary Pancreat Surg* 2004;**11**:79–83.
7. Konopke R, Kersting S, Distler M, Dietrich J, Gastmeier J, Heller A *et al.* Prognostic factors and evaluation of a clinical score for predicting survival after resection of colorectal liver metastases. *Liver Int* 2009;**29**:89–102.
8. Wimmer K, Schwarz C, Szabo C, Bodingbauer M, Tamandl D, Mittlbock M *et al.* Impact of neoadjuvant chemotherapy on clinical risk scores and survival in patients with colorectal liver metastases. *Ann Surg Oncol* 2017;**24**:236–243.
9. Ayez N, van der Stok EP, Grunhagen DJ, Rothbarth J, van Meerten E, Eggermont AM *et al.* The use of neo-adjuvant chemotherapy in patients with resectable colorectal liver metastases: clinical risk score as possible discriminator. *Eur J Surg Oncol* 2015;**41**:859–867.
10. Gomez D, Cameron IC. Prognostic scores for colorectal liver metastasis: clinically important or an academic exercise? *HPB (Oxford)* 2010;**12**:227–238.
11. Kanas GP, Taylor A, Primrose JN, Langeberg WJ, Kelsh MA, Mowat FS *et al.* Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors. *Clin Epidemiol* 2012;**4**:283–301.
12. Nordlinger B, Sorbye H, Glimelius B, Poston GJ, Schlag PM, Rougier P *et al.*; Fédération Francophone de Cancérologie Digestive (FFCD). Perioperative chemotherapy with FOLFOX4 and surgery *versus* surgery alone for resectable liver metastases from colorectal cancer (EORTC Intergroup trial 40983): a randomised controlled trial. *Lancet* 2008;**371**:1007–1016.
13. Nordlinger B, Sorbye H, Glimelius B, Poston GJ, Schlag PM, Rougier P *et al.*; Fédération Francophone de Cancérologie Digestive (FFCD). Perioperative FOLFOX4 chemotherapy and surgery *versus* surgery alone for resectable liver metastases from colorectal cancer (EORTC 40983): long-term results of a randomised, controlled, phase 3 trial. *Lancet Oncol* 2013;**14**:1208–1215.
14. Forrest LM, McMillan DC, McArdle CS, Angerson WJ, Dunlop DJ. Evaluation of cumulative prognostic scores based on the systemic inflammatory response in patients with inoperable non-small-cell lung cancer. *Br J Cancer* 2003;**89**:1028–1030.
15. McMillan DC. The systemic inflammation-based Glasgow Prognostic Score: a decade of experience in patients with cancer. *Cancer Treat Rev* 2013;**39**:534–540.
16. Shrotriya S, Walsh D, Bennani-Baiti N, Thomas S, Lorton C. C-reactive protein is an important biomarker for prognosis tumor recurrence and treatment response in adult solid tumors: a systematic review. *PLoS One* 2015;**10**:e0143080.
17. Lu X, Guo W, Xu W, Zhang X, Shi Z, Zheng L *et al.* Prognostic value of the Glasgow prognostic score in colorectal cancer: a meta-analysis of 9839 patients. *Cancer Manag Res* 2018;**11**:229–249.
18. Fruhling P, Hellberg K, Ejder P, Stromberg C, Urdzik J, Isaksson B. The prognostic value of C-reactive protein and albumin in patients undergoing resection of colorectal liver metastases. A retrospective cohort study. *HPB (Oxford)* 2021;**23**:970–978.
19. Forrest LM, McMillan DC, McArdle CS, Angerson WJ, Dunlop DJ. Comparison of an inflammation-based prognostic score (GPS) with performance status (ECOG) in patients receiving platinum-based chemotherapy for inoperable non-small-cell lung cancer. *Br J Cancer* 2004;**90**:1704–1706.
20. Forrest LM, McMillan DC, McArdle CS, Angerson WJ, Dagg K, Scott HR. A prospective longitudinal study of performance status, an inflammation-based score (GPS) and survival in patients with inoperable non-small-cell lung cancer. *Br J Cancer* 2005;**92**:1834–1836.
21. Brudvik KW, Jones RP, Giuliante F, Shindoh J, Passot G, Chung MH *et al. RAS* mutation clinical risk score to predict survival after resection of colorectal liver metastases. *Ann Surg* 2019;**269**:120–126.
22. Margonis GA, Sasaki K, Gholami S, Kim Y, Andreatos N, Rezaee N *et al.* Genetic and Morphological Evaluation (GAME) score for patients with colorectal liver metastases. *Br J Surg* 2018;**105**:1210–1220.
23. Lang H, Baumgart J, Heinrich S, Tripke V, Passalaqua M, Maderer A *et al.* Extended molecular profiling improves stratification and prediction of survival after resection of colorectal liver metastases. *Ann Surg* 2019;**270**:799–805.
24. Corcoran RB, Ebi H, Turke AB, Coffee EM, Nishino M, Cogdill AP *et al.* EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of *BRAF* mutant colorectal cancers to *RAF* inhibition with vemurafenib. *Cancer Discov* 2012;**2**:227–235.
25. Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D *et al.* ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol* 2016;**27**:1386–1422.
26. Tran B, Kopetz S, Tie J, Gibbs P, Jiang ZQ, Lieu CH *et al.* Impact of *BRAF* mutation and microsatellite instability on the pattern of metastatic spread and prognosis in metastatic colorectal cancer. *Cancer* 2011;**117**:4623–4632.
27. Meulendijks D, Henricks LM, Sonke GS, Deenen MJ, Froehlich TK, Amstutz U *et al.* Clinical relevance of DPYD variants c.1679T>G, c.1236G>A/HapB3, and c.1601G>A as predictors of severe fluoropyrimidine-associated toxicity: a systematic review and meta-analysis of individual patient data. *Lancet Oncol* 2015;**16**:1639–1650.
28. Sasaki K, Margonis GA, Andreatos N, Zhang XF, Buettner S, Wang J *et al.* The prognostic utility of the 'Tumor Burden Score' based on preoperative radiographic features of colorectal liver metastases. *J Surg Oncol* 2017;**116**:515–523.

29. Holch JW, Ricard I, Stintzing S, Modest DP, Heinemann V. The relevance of primary tumour location in patients with metastatic colorectal cancer: a meta-analysis of first-line clinical trials. *Eur J Cancer* 2017;**70**:87–98.
30. Kostner AH, Kersten C, Lowenmark T, Ydsten KA, Peltonen R, Isoniemi H *et al.* The prognostic role of systemic inflammation in patients undergoing resection of colorectal liver metastases: C-reactive protein (CRP) is a strong negative prognostic biomarker. *J Surg Oncol* 2016;**114**:895–899.
31. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996;**17**:343–346.
32. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT *et al.* The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012;**367**:1355–1360.
33. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;**338**:b2393.
34. He Y. Missing data analysis using multiple imputation: getting to the heart of the matter. *Circ Cardiovasc Qual Outcomes* 2010;**3**:98–105.
35. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* (2nd edn). Switzerland AG: Springer Nature, 2019.
36. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
37. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol* 2013;**13**:33.
38. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A *et al.*; PROGRESS Group. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* 2013;**346**:e5595.
39. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA *et al.*; PROGRESS Group. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013;**10**:e1001380.
40. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S *et al.*; PROGRESS Group. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;**10**:e1001381.
41. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW *et al.*; PROGRESS Group. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* 2013;**346**:e5793.
42. Van Cutsem E, Nordlinger B, Adam R, Kohne CH, Pozzo C, Poston G *et al.*; European Colorectal Metastases Treatment Group. Towards a pan-European consensus on the treatment of patients with colorectal liver metastases. *Eur J Cancer* 2006;**42**:2212–2221.
43. House MG, Ito H, Gonen M, Fong Y, Allen PJ, DeMatteo RP *et al.* Survival after hepatic resection for metastatic colorectal cancer: trends in outcomes for 1600 patients during two decades at a single institution. *J Am Coll Surg* 2010;**210**:744–752.
44. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;**338**:b606.
45. Ayez N, Lalmahomed ZS, van der Pool AE, Vergouwe Y, van Montfort K, de Jonge J *et al.* Is the clinical risk score for patients with colorectal liver metastases still useable in the era of effective neoadjuvant chemotherapy? *Ann Surg Oncol* 2011;**18**:2757–2763.
46. Sasaki K, Morioka D, Conci S, Margonis GA, Sawada Y, Ruzzenente A *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**:132–141.
47. Margonis GA, Andreatos N, Brennan MF. Predicting survival in colorectal liver metastasis: time for new approaches. *Ann Surg Oncol* 2020;**27**:4861–4863.
48. Shindoh J, Chun YS, Loyer EM, Vauthey JN. Non-size-based response criteria to preoperative chemotherapy in patients with colorectal liver metastases: the morphologic response criteria. *Curr Colorectal Cancer Rep* 2013;**9**:198–202.
49. Masi G, Vasile E, Loupakis F, Cupini S, Fornaro L, Baldi G *et al.* Randomized trial of two induction chemotherapy regimens in metastatic colorectal cancer: an updated analysis. *J Natl Cancer Inst* 2011;**103**:21–30.
50. Van Cutsem E, Kohne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A *et al.* Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009;**360**:1408–1417.
51. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N *et al.* Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;**21**:128–138.
52. Mallett S, Royston P, Waters R, Dutton S, Altman DG. Reporting performance of prognostic models in cancer: a review. *BMC Med* 2010;**8**:21.