Prognostic Nomogram for Liver Metastatic Colon Cancer Based on Histological Type, Tumor Differentiation, and Tumor Deposit: A TRIPOD Compliant Large-Scale Survival Study

Le Kuai¹,²†, Ying Zhang¹,²†, Ying Luo¹,², Wei Li³, Xiao-dong Li⁴,⁵, Hui-ping Zhang⁶, Tai-yi Liu⁶, Shuang-yi Yin³⁴ and Bin Li¹,²,⁷*

¹ Department of Dermatology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China, ² Institute of Dermatology, Shanghai Academy of Traditional Chinese Medicine, Shanghai, China, ³ Center for Translational Medicine, Huaize Hospital of Henan University, Kaifeng, China, ⁴ Department of Urology Surgery, Huaize Hospital of Henan University, Kaifeng, China, ⁵ Institute of Evidence-Based Medicine and Knowledge Translation, Henan University, Kaifeng, China, ⁶ Research and Development Center, Shanghai Applied Protein Technology Co., Ltd., Shanghai, China, ⁷ Shanghai Skin Disease Hospital, School of Medicine, Tongji University, Shanghai, China

Objective: A proportional hazard model was applied to develop a large-scale prognostic model and nomogram incorporating clinicopathological characteristics, histological type, tumor differentiation grade, and tumor deposit count to provide clinicians and patients diagnosed with colon cancer liver metastases (CLM) a more comprehensive and practical outcome measure.

Methods: Using the Transparent Reporting of multivariable prediction models for individual Prognosis or Diagnosis (TRIPOD) guidelines, this study identified 14,697 patients diagnosed with CLM from 1975 to 2017 in the Surveillance, Epidemiology, and End Results (SEER) 21 registry database. Patients were divided into a modeling group (n=9800), an internal validation group (n=4897) using computerized randomization. An independent external validation cohort (n=60) was obtained. Univariable and multivariate Cox analyses were performed to identify prognostic predictors for overall survival (OS). An independent external validation cohort (n=60) was obtained. Univariable and multivariate Cox analyses were performed to identify prognostic predictors for overall survival (OS). Subsequently, the nomogram was constructed, and the verification was undertaken by receiver operating curves (AUC) and calibration curves.

Results: Histological type, tumor differentiation grade, and tumor deposit count were independent prognostic predictors for CLM. The nomogram consisted of age, sex, primary site, T category, N category, metastasis of bone, brain or lung, surgery, and chemotherapy. The model achieved excellent prediction power on both internal (mean
AUC=0.811) and external validation (mean AUC=0.727), respectively, which were significantly higher than the American Joint Committee on Cancer (AJCC) TNM system. **Conclusion:** This study proposes a prognostic nomogram for predicting 1- and 2-year survival based on histopathological and population-based data of CLM patients developed using TRIPOD guidelines. Compared with the TNM stage, our nomogram has better consistency and calibration for predicting the OS of CLM patients.

**Keywords:** colon cancer, liver metastasis, database analysis, prognosis model, nomogram

1 INTRODUCTION

Globally, colon cancer is the third most common tumor and the second leading cause of cancer-related deaths (1, 2). The 5-year survival rate for colon cancer is 64.6%, while for synchronous metastasis, the patient’s survival is only 14.3%. Liver metastasis is the most frequently (17%) observed synchronous metastasis (3). It occurs in over 25% of patients initially and 50% of patients throughout the disease (4).

The prognosis of CLM varies significantly such that personalized prediction of CLM has become the focus of various studies, including those of the American Joint Committee on Cancer (AJCC) TNM system, which has been applied worldwide as the most authorized tool (5). However, the prediction accuracy of TNM staging is not satisfactory enough to predict outcomes (C-index=0.453) (6), which can relate to less predictors and classification on continuous variables (7).

Nomogram has been demonstrated to enhance predictive accuracy. Huang et al. developed two nomograms for the overall survival (OS) of patients with lung metastasis, having C-indexes of 0.754 and 0.749 (8). Furthermore, the C-index of a nomogram predicting the risk of bone metastasis has been reported to be 0.929 (9). Therefore, the present study will construct prognosis nomogram, providing clinicians with a more comprehensive outcome measure.

Recently, the identification of predictors was perceived great importance for prognosis nomogram. The histological type (10, 11), tumor differentiation grade (12), and tumor deposit count (13) have been recognized as independent predictors of colon cancer prognosis, as well as liver metastases. Several published studies have reported the prediction prognosis of histopathological predictors, respectively; however, insufficient sample sizes have limited the prognostic capabilities. Nevertheless, there have been a few large-scale CLM nomograms that have incorporated these histopathological indicators. The nomogram by Wu et al. ignored the influence of the histological type and the presence of synchronous metastasis and treatment information, which are important reasons for the nomogram’s low consistency (C-index=0.621).

Thus, based on data from the Surveillance, Epidemiology, and End Results (SEER) database, this study aims to develop a large-scale model and construct a nomogram that incorporates histological type, tumor differentiation grade, and tumor deposit count. TRIPOD guidelines were used for the development and verification process of the present study.

2 METHODS

This study design refers to the Prognosis Research Strategy (PROGRESS) (14). It uses the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) to demonstrate its research plan (15). The study followed the TRIPOD checklist (Supplementary Table 1). The flow diagram is displayed in **Figure 1**.

2.1 Participants

The data for this study were extracted from information recorded in the SEER 21 registry (16) from 1975 to 2017. The SEER database is a public cancer information registration database supported by the National Cancer Institute, collecting data from population registries to provide information on survival.
The analyzed data in this study are obtained from the SEER database, which is a multi-institution and multi-population registry. The inclusion criteria are as follows: (1) diagnosed with colon cancers (Site recode of ICD-O-3/WHO 2008: C180–C189; https://seer.cancer.gov/siterecode/icdo3_dwhoheme/index.html); (2) liver metastases; (3) complete survival data and follow-up data; (4) basic demographic information, sex and age; (5) histological information, histological type (ICD-O-3 Hist/behave), tumor differentiation grade (17), and tumor deposit (CS site-specific factor 4) (18); (6) clinical information, such as AJCC TNM category, bone/brain/lung metastasis (SEER Combined Mets at DX-bone/brain/lung), tumor size (CS tumor size), tumor location (primary site) (19), lymph (CS lymph nodes) (20), and CEA (CS site-specific factor 1) (21); (7) therapeutic measures that whether received surgery or surgery (Surgery Primary Site) (22) or chemotherapy. Missing entries (more than 80% of item missing or over 15 items were unknown) were excluded from predictive modeling analyses. For each variable, subtypes with frequency less than 5% were merged, to reduce small sample bias.

2.2 Ethics
Because the SEER data were de-identified, the study did not require either institutional review board approval or the subject’s informed consent. All procedures are in compliance with the 1964 Declaration of Helsinki and its subsequent amendments and standards.

2.3 Outcome
The outcome of the present study was death overtime, which was already assessed in the database. The definition of the outcome indicators and measurement methods was the same for all patients. Data collection of outcome indicators was performed from population registries of the SEER database, whose proponents did not participate in statistical analysis.

2.4 Predictors
In this study, information on the patient’s clinicopathological factors was collected and included histopathological indicators, such as tumor differentiation grade, tumor deposit, and histological type; population-based indicators, including age and sex; tumor-related indicators, including tumor site, tumor size, lymph, CEA, and metastasis, as well as treatment-related indicators of chemotherapy and surgery (surgery in primary site). The main elements of the TNM category was also included. Predictors were reliable and straightforward for clinicians. Age, sex, surgical history, and chemotherapy history were obtained by the clinical inquiry; tumor-related indicators such as tumor size, grade, tumor site, synchronous metastasis, etc., were obtained through previous disease assessment, which is approachable for clinicians. The predictors were defined according to the SEER classification (16; https://seer.cancer.gov/).

Prediction models are often adapted to diverse data for an improved prediction (23), and subpopulations with small sample sizes had been merged or removed to increase the accuracy.

2.5 Missing Data
Missing data was marked as “unknown” and was estimated for the prognosis model, allowing patients who lack some data to use this model. However, missing entries (more than 80% of item missing or over 15 items were unknown) were excluded from predictive modeling analyses.

2.6 Statistical Analysis
2.6.1 Modeling and Verifying Samples
A total of 14,757 patients were involved in this study. Patients were randomly divided into a modeling group (n=9800), internal verifying group (n=4897). An independent external verifying dataset (n=60) were obtained, which was completely independent from the data of model training, with varied demographic background characteristics. Clinical data were collected from medical record reviews in China.

2.6.2 Cox Regression Modeling
Univariate regression analysis was used to determine predictors with independent effects and to initially screen predictors. In addition, multivariate stepwise Cox regression was used to analyze statistically significant variables by hazard ratios (HR) and to predict the association of overall survival and related prognostic factors.

2.6.3 Nomogram Construction
A nomogram was constructed using R software after Cox regression analysis based on the modeling group. All statistical analyses were completed in R software (3rd edition 3.6.2; https://www.r-project.org).

2.7 Risk Groups
This study did not group the predicted probability into risk groups, as the prediction model is a basis for risk judgment.

2.8 Model Validation
Model internal verification was conducted using the SEER data sets, and the external validation was performed based on independent cohort. The consistency of the model was displayed using the C-index, the area under the receiver operating curves (AUCs). The model was acceptable when the AUC was over 0.7. The calibration curve was also used for model verification.

3 RESULTS
3.1 Participants
A total of 14,697 patients from SEER database were involved in this study. Patients were randomly divided into a modeling group (n=9800) and a verification group (n=4897). Figure 2 shows the baseline characteristics of all patients stratified by age; the median age of the modeling group and validation group was 65 years. Patients with CLM are more likely to be male (54.38%), regardless of race or region.

Adenocarcinoma seems to be related to liver metastasis; 13,434 (91.41%) patients exhibited adenocarcinoma in the
present study, while the remaining 1263 (8.59%) patients were divided into the group without adenocarcinoma for comparative analysis. Patients with a high degree of tumor differentiation grade are known to more likely metastasize; therefore, only 782 (5.32%) CLM patients presented as tumor differentiation grade I. In addition, our results demonstrated that tumor differentiation grade II (65.84%), or moderate differentiation, was present in the majority of CLM patients. Only 13.36% of patients presented positive in tumor deposit, which referred to a solitary tumor nodule that existed in the lymphatic drainage area of the primary tumor. Overall, 11,452 (77.92%) patients died, and the median survival was 14 months. Other population-based data are presented in Table 1.

3.2 Model Development

3.2.1 Cox Regression to Screen Predictors

Univariate regression analysis was carried out to predict the overall survival of CLM patients. The findings showed that only tumor size of 1-5cm, lung metastasis, and tumor deposit counts had a significant association with higher incidence of death (Table 2). Meanwhile, chemotherapy and surgery were also identified as significant predictors. Notably, the interaction between factors would influence the results of univariate regression. Therefore, we carried out further multivariable regression. Stepwise regression was used for the multivariable analysis, and the results showed that all predictors were found to be significantly associated with OS, except for the T category and tumor site (Table 3).

Based on these results, the Cox modeling in this study included twelve independent predictors (age, sex, histological type, tumor differentiation grade, tumor deposit, N category, bone metastasis, brain metastasis, lung metastasis, tumor size, surgery, and chemotherapy) of OS. As an essential indicator of the TNM system, the T category was included in the construction of the model. The predictive performance indicates that significant unknown factors remain (coded as 999, unknown or NOS in SEER database) for objective reasons, including data cannot be assessed, therefore, they may be of particular clinical significance (24).

3.2.2 Prognosis Model Development

Taking 0.05 as threshold of P-value, we used the hazard ratio (HR) values and coefficient of the multivariable regression model to represent the parameters of the independent variables for death. We identified a predictor as a risk factor for death when the corresponding coefficient was >0, or the HR value was >1 significantly. Among these clinical features, tumor differentiation grade IV was the most important factor for prognosis (β=0.544, HR=1.720), which indicated that the risk of death in grade IV was 1.720 times that of grade II followed by tumor differentiation grade III (β=0.385, HR=1.470). Conversely, tumor differentiation grade I (β=0.408, HR=0.665) was considered a protective factor for CLM OS. In different histological types, not exhibiting adenocarcinoma (β=0.160, HR=1.170) resulted in a poorer prognosis, the risk of which was 1.17 times that of adenocarcinoma patients. In addition, the occurrence of tumor deposit counts (β=0.112, HR=1.120) also increased the risk of death by 1.12 times.

As for synchronous metastasis, the prognosis of patients with brain metastasis (β=0.466, HR=1.540) was worse, and the impact on prognosis was more incredible than bone metastasis (β=0.562, HR=1.440) or lung metastasis (β=0.278, HR=1.320).

Although the influence of sex is not apparent, it can be concluded from the model that the risk of death for men was 1.09 times that of women. The risk of death for patients who received surgery (β=−0.806, HR=0.447) was only 45% of those who did not, and the chemotherapy seems to be more effective in improving OS (β=−1.086, HR=0.338).

3.2.3 Nomogram Construction

Based on the multivariable regression model results, specific scores of each predictor and total scores were plotted (Figure 3). The score predictors were as follows: male, 2.5; not adenocarcinoma, 5; tumor differentiation grade II, 12.5; tumor differentiation grade III, 22.5; IV, 27.5; tumor deposit, 2.5; bone metastasis, 10; brain metastasis, 12.5; lung metastasis, 7.5; no surgery, 25; and no chemotherapy, 30. Applying this model for clinical prognosis prediction is quite convenient. Consider a female patient aged 70 years old, having a 4-cm-long tumor in
## TABLE 1 | The baseline characteristics of enrolled CLM patients.

|                        | All samples | Modeling | Validation |
|------------------------|-------------|----------|------------|
|                        | N           | Percentage (%) | N           | Percentage (%) | N           | Percentage (%) |
| Sex                    |             |         |             |            |             |            |
| Female                 | 6705        | 45.62   | 4443        | 45.34      | 2262        | 46.19       |
| Male                   | 7992        | 54.38   | 5357        | 54.66      | 2635        | 53.81       |
| Tumor site             |             |         |             |            |             |            |
| Colon, unknown         | 21          | 0.14    | 16          | 0.16       | 5           | 0.10        |
| Left colon             | 6135        | 41.74   | 4067        | 41.50      | 2068        | 42.23       |
| Overlapping lesion     | 234         | 1.59    | 154         | 1.57       | 80          | 1.63        |
| Right colon            | 7041        | 47.91   | 4706        | 48.02      | 2335        | 47.68       |
| Transverse colon       | 1266        | 8.61    | 857         | 8.74       | 409         | 8.35        |
| Histological type      |             |         |             |            |             |            |
| Adenocarcinoma         | 13434       | 91.41   | 9007        | 91.91      | 4427        | 90.40       |
| Not adenocarcinoma     | 1263        | 8.59    | 793         | 8.09       | 470         | 9.60        |
| Tumor differentiation grade |       |         |             |            |             |            |
| II                     | 9677        | 65.84   | 6470        | 66.02      | 3207        | 65.49       |
| IV                     | 738         | 5.02    | 476         | 4.86       | 262         | 5.35        |
| I                      | 782         | 5.32    | 507         | 5.17       | 275         | 5.62        |
| III                    | 3500        | 23.81   | 2347        | 23.95      | 1153        | 23.55       |
| T category             |             |         |             |            |             |            |
| T0                     | 3           | 0.02    | 2           | 0.02       | 1           | 0.02        |
| T1                     | 1171        | 7.97    | 785         | 8.01       | 386         | 7.88        |
| T2                     | 354         | 2.41    | 249         | 2.54       | 105         | 2.14        |
| T3                     | 6055        | 41.20   | 4048        | 41.31      | 2007        | 40.88       |
| T4                     | 4626        | 31.48   | 3077        | 31.40      | 1549        | 31.83       |
| TX                     | 2488        | 16.93   | 1639        | 16.72      | 849         | 17.34       |
| N category             |             |         |             |            |             |            |
| N0                     | 3872        | 26.35   | 2580        | 26.33      | 1292        | 26.38       |
| N1                     | 5104        | 34.73   | 3480        | 35.51      | 1696        | 34.63       |
| N2                     | 4740        | 32.25   | 3157        | 32.21      | 1583        | 32.33       |
| NX                     | 981         | 6.67    | 655         | 6.68       | 326         | 6.66        |
| Bone metastasis        |             |         |             |            |             |            |
| No                     | 14151       | 96.28   | 9423        | 96.15      | 4728        | 96.55       |
| Yes                    | 546         | 3.72    | 377         | 3.85       | 169         | 3.45        |
| Brain metastasis       |             |         |             |            |             |            |
| No                     | 14589       | 99.27   | 9728        | 99.27      | 4861        | 99.26       |
| Yes                    | 108         | 0.73    | 72          | 0.73       | 36          | 0.74        |
| Lung metastasis        |             |         |             |            |             |            |
| No                     | 12017       | 81.76   | 8002        | 81.65      | 4015        | 81.99       |
| Yes                    | 2680        | 18.24   | 1798        | 18.35      | 882         | 18.01       |
| Tumor deposit          |             |         |             |            |             |            |
| No                     | 6127        | 41.69   | 4088        | 41.51      | 2059        | 42.05       |
| Unknown                | 6607        | 44.95   | 4409        | 44.99      | 2198        | 44.88       |
| Yes                    | 1963        | 13.36   | 1323        | 13.50      | 640         | 13.57       |
| Tumor size             |             |         |             |            |             |            |
| 0-1cm                  | 11664       | 79.36   | 7789        | 79.48      | 3875        | 79.13       |
| 1-5cm                  | 32          | 0.22    | 23          | 0.23       | 9           | 0.18        |
| Unknown                | 3001        | 20.42   | 1988        | 20.29      | 1013        | 20.69       |
| Lymph                  |             |         |             |            |             |            |
| No                     | 13445       | 91.48   | 8948        | 91.31      | 4497        | 91.83       |
| Unknown                | 1232        | 8.52    | 838         | 8.55       | 394         | 8.05        |
| Yes                    | 20          | 0.14    | 14          | 0.14       | 6           | 0.12        |
| CEA                    |             |         |             |            |             |            |
| Negative               | 1371        | 9.33    | 905         | 9.23       | 466         | 9.52        |
| Positive               | 8553        | 58.20   | 5725        | 58.42      | 2828        | 57.75       |
| Unknown                | 4773        | 32.48   | 3170        | 32.35      | 1603        | 32.73       |
| Surgery                |             |         |             |            |             |            |
| No                     | 4274        | 29.08   | 2845        | 29.03      | 1429        | 29.18       |
| Yes                    | 10423       | 70.92   | 6955        | 70.97      | 3468        | 70.82       |
| Chemotherapy           |             |         |             |            |             |            |
| No/Unknown             | 4980        | 33.88   | 3311        | 33.79      | 1669        | 34.08       |
| Yes                    | 9717        | 66.12   | 6489        | 66.21      | 3228        | 65.92       |

P-value was adopted on comparison of modeling and verifying dataset. CLM, colon cancer liver metastases; CEA, carcinoembryonic antigen.
| Variable                        | β     | HR  | se(β) | Z     | P      | 95% CI (HR) |
|--------------------------------|-------|-----|-------|-------|--------|-------------|
| Age                            | -0.003| 0.997| 0.002 | -1.751| 0.0799 | 0.994       | 1.000      |
| Sex                            |       |     |       |       |        |             |            |
| Female                         |       |     |       |       |        |             |            |
| Male                           | 0.015 | 1.015| 0.043 | 0.355 | 0.7229 | 0.933       | 1.105      |
| Tumor site                     |       |     |       |       |        |             |            |
| Colon, unknown                 |       |     |       |       |        |             |            |
| Colon                           |       |     |       |       |        |             |            |
| Left colon                     | -0.730| 0.482| 0.253 | -2.883| 0.0039 | 0.293       | 0.792      |
| Overlapping lesion             | -0.554| 0.575| 0.305 | -1.818| 0.0691 | 0.317       | 1.044      |
| Right colon                    | -0.688| 0.503| 0.254 | -2.708| 0.0068 | 0.306       | 0.827      |
| Transverse colon               | -0.626| 0.535| 0.263 | -2.381| 0.0173 | 0.319       | 0.895      |
| Histological type              |       |     |       |       |        |             |            |
| Adenocarcinoma                 |       |     |       |       |        |             |            |
| Not adenocarcinoma             | 0.047 | 1.048| 0.087 | 0.545 | 0.5858 | 0.885       | 1.242      |
| Tumor differentiation grade    |       |     |       |       |        |             |            |
| II                             |       |     |       |       |        |             |            |
| IV                             | -0.028| 0.973| 0.136 | -0.203| 0.8391 | 0.745       | 1.270      |
| I                              | -0.001| 0.999| 0.084 | -0.012| 0.9906 | 0.848       | 1.178      |
| III                            | -0.029| 0.972| 0.061 | -0.473| 0.6364 | 0.826       | 1.095      |
| T category                     |       |     |       |       |        |             |            |
| TX1                            | 1.448 | 4.256| 1.006 | 1.440 | 0.1498 | 0.593       | 30.546     |
| T2                             | 1.421 | 4.140| 1.006 | 1.413 | 0.1578 | 0.577       | 29.726     |
| T3                             | 1.267 | 3.550| 1.002 | 1.285 | 0.2060 | 0.498       | 25.290     |
| T4                             | 1.245 | 3.474| 1.002 | 1.242 | 0.2142 | 0.487       | 24.776     |
| TX1                            | 1.268 | 5.091| 1.004 | 1.621 | 0.1050 | 0.712       | 36.421     |
| N category                     |       |     |       |       |        |             |            |
| N1                             | -0.012| 0.981| 0.052 | -0.376| 0.7073 | 0.886       | 1.086      |
| N2                             | -0.164| 0.849| 0.057 | -2.858| 0.0043 | 0.759       | 0.950      |
| NX                             | 0.061 | 1.063| 0.129 | 0.474 | 0.6358 | 0.826       | 1.368      |
| Bone metastasis                |       |     |       |       |        |             |            |
| No                             |       |     |       |       |        |             |            |
| Yes                            | -0.089| 0.915| 0.187 | -0.476| 0.6340 | 0.634       | 1.320      |
| Brain metastasis               |       |     |       |       |        |             |            |
| No                             |       |     |       |       |        |             |            |
| Yes                            | 0.871 | 2.390| 0.449 | 1.942 | 0.0521 | 0.992       | 5.758      |
| Lung metastasis                |       |     |       |       |        |             |            |
| No                             |       |     |       |       |        |             |            |
| Yes                            | 0.225 | 1.252| 0.066 | 3.393 | 0.0007 | 1.100       | 1.426      |
| Tumor deposit                  |       |     |       |       |        |             |            |
| No                             |       |     |       |       |        |             |            |
| Unknown                        | 0.112 | 1.119| 0.050 | 2.260 | 0.0238 | 1.015       | 1.233      |
| Yes                            | 0.386 | 1.472| 0.061 | 6.307 | 0.0000 | 1.305       | 1.660      |
| Tumor size                     |       |     |       |       |        |             |            |
| 0-1cm                          |       |     |       |       |        |             |            |
| 1-5 cm                         | 1.064 | 2.897| 0.501 | 2.123 | 0.0338 | 1.085       | 7.735      |
| unknown                        | 0.132 | 1.141| 0.061 | 2.151 | 0.0315 | 1.012       | 1.286      |
| Lymph                          |       |     |       |       |        |             |            |
| No                             |       |     |       |       |        |             |            |
| Unknown                        | 0.143 | 1.153| 0.104 | 1.376 | 0.1689 | 0.941       | 1.413      |
| Yes                            | 0.968 | 2.634| 0.578 | 1.676 | 0.0938 | 0.849       | 8.176      |
| CEA                            |       |     |       |       |        |             |            |
| Negative                       |       |     |       |       |        |             |            |
| Positive                       | -0.050| 0.951| 0.075 | -0.669| 0.5038 | 0.822       | 1.101      |
| Unknown                        | 0.025 | 1.025| 0.079 | 0.313 | 0.7541 | 0.878       | 1.196      |
| Surgery                        |       |     |       |       |        |             |            |
| No                             |       |     |       |       |        |             |            |
| Yes                            | -0.357| 0.700| 0.059 | -6.052| 0.0000 | 0.623       | 0.785      |
| Chemotherapy                   |       |     |       |       |        |             |            |
| No                             |       |     |       |       |        |             |            |
| Yes                            | -0.025| 0.976| 0.061 | -0.407| 0.6841 | 0.867       | 1.099      |

OS, overall survival; HR, hazard ratio; CI, confidence interval.
the left colon, tumor differentiation grade II, and histological type of adenocarcinoma, without tumor deposit; grade T2, N1, and lung metastasis, who received chemotherapy, but no surgery. Her total score is 174, and clinicians could draw a vertical line at the corresponding position to obtain the corresponding predicted survival rate. Clinicians can thus use the visual nomogram to directly calculate the scores of the indicators and get the survival rate of patients to formulate suitable individual treatments.

3.3 Model Validation
The predicted distributions of the score, 1-year survival, and 2-year survival in the modeling and internal validation groups were consistent (Figure 4). A calibration plot of the 1-year OS
demonstrated good calibration between prediction and actual survival (Figure 5). The C-indexes of the modeling group and internal verification group were 0.751 and 0.752 (Table 4). The ROC curve analysis (Figures 6, 7) was used for nomogram validation, and the value of the area under the curve (AUC) represented the consistency of the model. The results showed that the AUC of the modeling group was 0.822 at 1-year and 0.801 at 2-year, and in the internal verification group, the AUC was 0.822 for 1-year and 0.799 for the 2-year survival (mean AUC=0.811). The ROC analysis of external verification (Figure 7 and Supplementary Table 2) indicated a good predictive performance of the model, in which the AUC values for predicting 1- and 2-year OS were 0.786 and 0.667 respectively. Albeit on small samples, this nomogram showed reliable predictive power in the external discrimination test (mean AUC=0.727).
FIGURE 5 | Calibration plots of the nomogram for 1-year and 2-year prediction for OS of CLM patients. OS, overall survival; CLM, colon cancer liver metastases.

TABLE 4 | Multivariable regression for OS.

| Data set   | C     | Dxy  | aDxy | SD   | Z    | P  | n  |
|------------|-------|------|------|------|------|----|----|
| Modeling   | 0.751 | -0.502 | 0.502 | 0.006 | 89.56 | 0  | 9800 |
| Validation | 0.752 | -0.498 | 0.498 | 0.008 | 62.41 | 0  | 4897 |

FIGURE 6 | ROC of the nomogram for 1-year and 2-year internal prediction. ROC, receiver operating curves, AUC, area under the curve; FP, false positive; TP, true positive.
4 DISCUSSION

In this study, we constructed a predictive model for CLM patients using multi-institutional clinical data from the SEER database. The prognosis nomogram for OS was established and its consistency and calibration were verified.

The predictors included in the prognostic nomogram were age, sex, tumor site, histological type, tumor differentiation grade, T category, N category, bone metastasis, lung metastasis, brain metastasis, tumor deposit, surgery, and chemotherapy.

Multivariate regression analysis indicated that age was an essential prognostic risk factor. Several models have shown that age is a risk factor (25, 26), illustrating that the elderly have a poorer prognosis, especially when over 60 years old (6). Our nomogram demonstrated that tumor differentiation grade IV, undifferentiated, has a poorer prognosis than tumor differentiation grade I, well-differentiated. Consistent with this result, the tumor differentiation grade was observed to be associated with the faster growth rate and more aggressive ability of tumor cells, leading to distant metastasis of CLM (27) and OS. Tumor deposits were found to be a statistically significant factor, which may be considered as a third mixed pathway of tumor migration and invasion in perivascular, perineural, or mesentery (28).

What differed from previous studies was that T0 and T1 were identified as risk factors for CLM, which may be related to the highly malignant characteristics of the tumors in the early stages (9). Adenocarcinomas, such as mucinous adenocarcinoma or signet ring cell carcinoma, are more likely to undergo peritoneal metastasis (29), always leading to poorer prognosis. Still, early mucinous adenocarcinoma may have a better prognosis. However, our results indicate that patients with adenocarcinoma will have a better prognosis, which may be related to the classification of the histological subtypes, but this needs further research to be validated.

Validation of the nomogram resulted in a reliable ability to discriminate events. The AUC has been generally accepted and widely used for model validation. The model achieved excellent prediction power on both internal (mean AUC=0.811) and external validation (mean AUC=0.727), respectively, which were substantially superior to the American Joint Committee on Cancer (AJCC) TNM system. This study will provide clinicians and patients diagnosed with colon cancer liver metastases (CLM) a more comprehensive and practical outcome measure, to help clinicians assess patient prognosis and determine personalized treatment decisions.

This study developed and validated a prognosis model to predict the survival of CLM patients under the guidance of both the PROGRESS framework and the TRIPOD prediction research report for the first time. Previously developed nomograms (30-33) could be more convincing if they expanded their sample size. We included as much population as possible from the SEER database, covering 35% population of the US. Several predictive models of CLM had been published to predict the survival rate of patients after hepatectomy (34, 35), while patients who were suitable for hepatectomy accounted for only 6.1% of patients (36), which led to the limitation of the model in clinical use. Wu Q et al. developed nomograms of CLM from the SEER database but excluded histological type, and synchronous metastasis, without any treatment-related data. This study incorporated these significant predictors, including more available clinical indicators as predictors to improve model applicability.

Internal and external validation of the current nomogram demonstrated high accuracy in predicting OS of CLM patients. However, three limitations must be resolved. The first problem relates to the inevitable lack of other clinicopathological factors of the SEER data set, some vital prognostic factors of liver metastasis, such as the type of liver resection, should be included in future research. Secondly, model focusing on subtypes of adenocarcinoma, or primary surgery type should be included in furtherer, using specific data set. Furthermore, although the external verification indicated an excellent predictive effect, the AUC value of 1-year OS prediction was slightly lower than 2-year, leading requirement to better external validation using large samples of independent multicenter cohorts.

5 CONCLUSION

In conclusion, based on the TRIPOD prediction research report, we established and validated nomograms to predict the 1- and 2-year survival based on histopathological and population-based data of CLM patients. Compared with the traditional staging
system TNM, our nomogram achieved relatively good discrimination and calibration.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

**AUTHOR CONTRIBUTIONS**

LK and YZ conceived and designed the study and provided administrative support. YL, WL, and X-DL collected and assembled the data. H-PZ and T-YL performed the external validation. S-YY and BL analyzed and interpreted the data. All authors contributed to the article and approved the submitted version.

**FUNDING**

This study was supported by the National Key Research and Development Program of China (No. 2018YFC1705305), the NSFC of China (No. 81973860, 81904214, 82004235), the Shanghai Development Office of TCM [No. ZY(2018-2020)-FWTX-4010, ZY(2018-2020)-FWTX-1008], and Dermatology Department of Traditional Chinese Medicine, Clinical Key Specialty Construction Project of Shanghai (No. shslczdzk05001), Xinglin Youth Scholar of Shanghai University of Traditional Chinese Medicine (no. RY411.33.10), Shanghai Sailing Program (no. 21YF1448100), and Peiran Scholar of Shanghai University of Traditional Chinese Medicine (no. 2015811038).

**ACKNOWLEDGMENTS**

We thank Charlesworth Author Services for English language editing.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.604882/full#supplementary-material

**REFERENCES**

1. Arnold M, Sierra MS, Lavranos S, Soerjomataram I, Jemal A, Bray F. Global Patterns and Trends in Colon Cancer Incidence and Mortality. *Gut* (2017) 66(4):683–91. doi: 10.1136/gutjnl-2015-310912
2. Allemani C, Weir HK, Carreira H, Farewell V, Spika D, Wang XS, et al. Global Surveillance of Cancer Survival 1995-2009: Analysis of Individual Data for 25,676,887 Patients From 279 Population-Based Registries in 67 Countries (CONCORD-2). *Lancet* (2015) 385(9972):977–1010. doi: 10.1016/S0140-6736(14)62038-9
3. Geest LGM, Lam BJ, Koopman M, Verhoef C, Elderink MAG, Wilt JHW. Nationwide Trends in Incidence, Treatment, and Survival of Colon Cancer Patients With Synchronous Metastases. *Clin Exp Metastasis* (2015) 32(5):457–65. doi: 10.1007/s10438-015-9719-0
4. Neal CP, Nana GR, Jones M, Cairns V, Wee N, Isherwood J, et al. Repeat Hepatectomy Is Independently Associated With Favorable Long-Term Outcome in Patients With Colon Liver Metastases. *Cancer Med* (2017) 6(2):331–8. doi: 10.1002/cam4.872
5. Weiser MR. AJCC 8th Edition: Colon Cancer. *Ann Surg Oncol* (2018) 25(6):1454–5. doi: 10.1245/s10434-018-6462-1
6. Wu Q, Wang WJ, Huang YQ, Fang SY, Guan YJ. Nomograms for Estimating Survival in Patients With Liver-Only Colon Metastases: A Retrospective Study. *Int J Surg* (2018) 60:1–8. doi: 10.1016/j.ijsu.2018.10.032
7. Pei JP, Zhang CD, Liang Y, Zhang C, Wu KZ, Zhao ZM, et al. Novel Nomograms Individually Predicting Overall Survival of Non-Metastatic Colon Cancer Patients. *Front Oncol* (2020) 10:733. doi: 10.3389/fonc.2020.00733
8. Huang WY, Zhao MN, Yin JC, Lu T, Yang XD, Yuan GD, et al. Pulmonary Metastasis in Newly Diagnosed Colon-Rectal Cancer: A Population-Based Nomogram Study. *Int J Colon Dis* (2019) 34(5):867–78. doi: 10.1007/s00384-019-03270-w
9. Han LY, Dai WX, Mo SB, Xiang WQ, Li QG, Xu Y, et al. Nomogram to Predict the Risk and Survival of Synchronous Bone Metastasis in Colon Cancer: A Population-Based Real-World Analysis. *Int J Colon Dis* (2020) 35(8):1575–85. doi: 10.1007/s00384-020-03612-z
10. Huang Y, Ge KX, Fu GS, Chu JF, Wei W. Mucinous Histological Might Be an Indicator for Enhanced Survival Benefit of Chemotherapy in Stage II Colon Cancer. *Front Med (Lausanne)* (2020) 7:205. doi: 10.3389/fmed.2020.00205
11. Powers BD, Felder SI, Izzard R, Dessureault S, Dineen SP. The Impact of Histologic Subtype on Receipt of Adjuvant Chemotherapy and Overall Survival in Stage III Colon Cancer: A Retrospective Cohort Analysis. *J Gastrointest Cancer* (2021) 52:719–27. doi: 10.1007/s12029-020-00460-6
12. Tsai MS, Su YH, Ho MC, Lai SP, Chen TP, Lai HS, et al. Clinicopathological Features and Prognosis in Resectable Synchronous and Metachronous Colon Liver Metastases. *Ann Surg Oncol* (2007) 14(2):786–94. doi: 10.1245/s10434-006-9215-9
13. D’Souza N, Shaw A, Lord A, Balyasnikova S, Abulafi M, Tekkis P, et al. Assessing a Staging System for Sigmoid Colon Cancer Based on Tumor Deposits and Extramural Venous Invasion on Computed Tomography. *JAMA Netw Open* (2019) 2(12):e1916987. doi: 10.1001/jamanetworkopen.2019.16987
14. Steyerberg EW, Moons KGM, Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. *PloS Med* (2013) 10(2):e1001381. doi: 10.1371/journal.pmed.1001381
15. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *BMJ* (2015) 350:g7594. doi: 10.1136/bmj.g7594
16. Howlard N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, et al. SEER Cancer Statistics Review, 1975-2017. Bethesda, MD: National Cancer Institute (2020). Available at: https://seer.cancer.gov/csr/1975_2017/. based on November 2019 SEER data submission, posted to the SEER web site, April 2020.
17. Alexander D, Ihalo N, Chatla C, Steinhauser J, Funkhouser E, Coffer ES, et al. High-Grade Tumor Differentiation Is an Indicator of Poor Prognosis in African Americans With Colonic Adenocarcinomas. *Cancer* (2005) 103:2163–70. doi: 10.1002/cncr.21021
18. Mirkin KA, Kulaylat AS, Holfbein BS, Messaris E. Prognostic Significance of Tumor Deposits in Stage III Colon Cancer. *Ann Surg Oncol* (2018) 25:3179–85. doi: 10.1245/s10434-018-6661-9
19. Shen H, Yang J, Huang Q, Jiang M-J, Tan Y-N, Fu J-F, et al. Different Treatment Strategies and Molecular Features Between Right-Sided and Left-Sided Colon Cancers. *World J Gastroenterol* (2015) 21:6470–8. doi: 10.3748/wjg.v21.i21.6470
20. Parsons HM, Tuttle TM, Kuntz KM, Begun JW, McGovern PM, Virmig BA. Association Between Lymph Node Evaluation for Colon Cancer and Node Positivity Over the Past 20 Years. JAMA (2011) 306:1089–97. doi: 10.1001/
jama.2011.1285
21. Bhatti I, Patel M, Dennison AR, Thomas MW, Garcea G. Utility of Postoperative CEA for Surveillance of Recurrence After Resection of Primary Colorectal Cancer. Int J Surg (2015) 16:123–8. doi: 10.1016/j.ijsu.2015.03.002
22. Xu J, Ma T, Ye Y, Pan Z, Lu D, Pan F, et al. Surgery on Primary Tumor Shows Survival Benefit in Selected Stage IV Colon Cancer Patients: A Real-World Study Based on SEER Database. J Cancer (2020) 11:3567–79. doi: 10.7150/jca.45318
23. Kadam NN, Jagadish SVK, Struik PC, van der Linden CG, Yin X. Incorporating Genome-Wide Association Into Eco-Physiological Simulation to Identify Markers for Improving Rice Yields. J Exp Bot (2019) 70(9):2573–86. doi: 10.1093/jxb/erz120
24. Dasari A, Shen C, Halperin D, Zhao B, Zhou S, Xu Y, et al. Trends in the Incidence, Prevalence, and Survival Outcomes in Patients With Neuroendocrine Tumors in the United States. JAMA Oncol (2017) 3:1335–42. doi: 10.1001/jamaoncol.2017.0589
25. Ngu JCY, Kuo LJ, Teo NZ. Minimally Invasive Surgery in the Geriatric Patient With Colon Cancer. J Gastrointest Oncol (2020) 11(3):540–4. doi: 10.21037/jgo.2020.02.02
26. Shida D, Inoue M, Tanabe T, Moritani K, Tsukamoto S, Yamauchi S, et al. Prognostic Impact of Primary Tumor Location in Stage III Colorectal Cancer: Right-Sided Colon Versus Left-Sided Colon Versus Rectum: A Nationwide Multicenter Retrospective Study. J Gastroenterol (2020) 55(10):958–68. doi: 10.1007/s00535-020-01706-7
27. Serra PS, Pallisera LA, Mora LL, Granados MA, Gallardo S, Del PZC, et al. Multidisciplinary Management and Optimization of Frail or High Surgical Risk Patients in Colorectal Cancer Surgery: Prospective Observational Analysis. Cir Esp (2020) 98(7):389–94. doi: 10.1016/j.ciresp.2020
28. Awad J, Armin E, Audry SP, Dominic V, Michael V, David L, et al. An Up-To-Date Predictive Model for Rectal Cancer Survival: Reflecting Tumor Biology and Clinical Factors. Am J Surg (2019) 219(5):515–20. doi: 10.1016/j.amjsurg.2019.10.036
29. Wang R, Ma X, Li Y, He Y, Huang D, Cai S, et al. The Characteristics and Prognostic Effect of E-Cadherin Expression in Colorectal Signet Ring Cell Carcinoma. PLoS One (2016) 11(8):e0160527. doi: 10.1371/journal.pone.0160527
30. Li MY, Li XY, Guo Y, Miao Z, Liu XM, Guo SX, et al. Development and Assessment of an Individualized Nomogram to Predict Colorectal Cancer Liver Metastases. Quant Imaging Med Surg (2020) 10(2):397–414. doi: 10.21037/qims.2019.12.16
31. Higuchi A, Aoyama T, Kazama K, Murakawa M, Atsumi Y, Katayama Y, et al. Beppu’s Nomogram Score Is an Independent Prognostic Factor for Colorectal Liver Metastasis Receiving Perioperative Chemotherapy and/or Targeted Therapy. In Vivo (2019) 33(4):1301–6. doi: 10.21873/invivo.11603
32. Wang Y, Zheng JS, Chen H, Hu CX, Sun B, Wang HY, et al. A Prognostic Nomogram for Colorectal Cancer Liver Metastases After Percutaneous Thermal Ablation. Int J Hyperthermia (2018) 34(6):853–62. doi: 10.1080/02656736.2017.1368095
33. Cheng J, Wei JW, Tong T, Sheng WQ, Zhang YL, Han YQ, et al. Prediction of Histopathologic Growth Patterns of Colorectal Liver Metastases With a Noninvasive Imaging Method. Ann Surg Oncol (2019) 26(13):4587–98. doi: 10.1245/s10434-019-07910-x
34. Kattan MW, Gönen M, Jarnagin WR, DeMatteo R, D’Angelica M, Martin W, et al. A Nomogram for Predicting Disease-Specific Survival After Hepatic Resection for Metastatic Colorectal Cancer. Ann Surg (2008) 247(2):282–7. doi: 10.1097/SLA.0b013e31815ed67b
35. Matias M, Casa NM, Faria M, Pires R, Tato CJ, Ribeiro L, et al. Prognostic Factors After Liver Resection for Colorectal Liver Metastasis. Acta Med Port (2015) 28(3):357–69. doi: 10.20344/amp.4816
36. Cummings LC, Payes JD, Cooper GS. Survival After Hepatic Resection in Metastatic Colorectal Cancer: A Population-Based Study. Cancer (2007) 109(4):718–26. doi: 10.1002/cncr.22448

Conflict of Interest: Authors H-PZ and T-YL were employed by Research and Development Center, Shanghai Applied Protein Technology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Kuai, Zhang, Luo, Li, Li, Zhang, Liu, Yin and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.