A Routine Laboratory Data–Based Model for Predicting Recurrence After Curative Resection of Stage II Colorectal Cancer

Zhong Ye, MD, PhD\(^a\); Chun Wang, MD, PhD\(^a\); Limin Guo, MD\(^a\); Juan P. Palazzo, MD\(^a\); Zhixing Han, MD\(^b\); Yinzhi Lai, PhD\(^b\); Jing Jiang, MPH\(^b\); James A. Posey, MD\(^b\); Atrayee Basu Mallick, MD\(^b\); Bingshan Li, PhD\(^b\); Li Jiang, MD\(^c\); and Hushan Yang, PhD\(^a\)

**Abstract**

**Background:** Use of chemotherapy in stage II colorectal cancer (CRC) is controversial because it improves survival only in some patients. We aimed to develop a statistical model using routine and readily available blood tests to predict the prognosis of patients with stage II CRC and to identify which patients are likely to benefit from chemotherapy. **Methods:** We divided 422 patients with stage II CRC into a training and a testing set. The association of routine laboratory variables and disease-free survival (DFS) was analyzed. A prognostic model was developed incorporating clinically relevant laboratory variables with demographic and tumor characteristics. A prognostic score was derived by calculating the sum of each variable weighted by its regression coefficient in the model. Model performance was evaluated by constructing receiver operating characteristic curves and calculating the area under the curve (AUC). **Results:** Significant associations were seen between 5 laboratory variables and patient DFS in univariate analyses. After stepwise selection, 3 variables (carcinoembryonic antigen, hemoglobin, creatinine) were retained in the multivariate model with an AUC of 0.75. Compared with patients with a low prognostic score, those with a medium and high prognostic score had a 1.99- and 4.78-fold increased risk of recurrence, respectively. The results from the training set were validated in the testing set. Moreover, chemotherapy significantly improved DFS in high-risk patients, but not in low- and medium-risk patients. **Conclusions:** A routine laboratory variable–based model may help predict DFS of patients with stage II CRC and identify high-risk patients more likely to benefit from chemotherapy.

Colorectal cancer (CRC) is highly treatable by surgical removal if diagnosed at an early stage, and the 5-year survival rate for localized CRC is estimated to be 90%.\(^1\) Adjuvant chemotherapy has been clinically proven to reduce CRC recurrence, and its use is mainly determined by tumor stage in the current clinical setting. Patients with stage I CRC generally do not receive chemotherapy because most of them are cured by surgery alone. In comparison, those with stage III CRC receive chemotherapy as standard of care given the high

\(^{a}\)Department of Medical Oncology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, Pennsylvania; \(^{b}\)Department of Environmental Health, School of Public Health, Nan tong University, Nan tong, China; \(^{c}\)Department of Hepatobiliary Surgery, Beijing Ditan Hospital, Capital Medical University, Beijing, China; \(^{d}\)Department of Pathology, Thomas Jefferson University, Philadelphia, Pennsylvania; \(^{e}\)Department of Urology, Beijing Ditan Hospital, Capital Medical University, Beijing, China; and \(^{f}\)Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee.

\(^*\)These authors contributed equally.

\(^†\)Co-last authors.

Submitted August 14, 2017; accepted for publication May 10, 2018. The authors have disclosed that they have no financial interests, arrangements, affiliations, or commercial interests with the manufacturers of any products discussed in this article or their competitors. This work was supported by grants from NCI (CA162201); the Pennsylvania Department of Health and Thomas Jefferson University; and the American Cancer Society (123741-RSG-13-003-01-CCE).


**Correspondence:** Hushan Yang, PhD, Division of Population Science, Department of Medical Oncology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA 19107. Email: hushan.yang@jefferson.edu; and Li Jiang, MD, Department of Hepatobiliary Surgery, Beijing Ditan Hospital, Capital Medical University, Beijing 100015, China. Email: lijiang1963@163.com
recurrence rate. However, whether adjuvant chemotherapy should be used in patients with stage II CRC is a matter of ongoing debate. Indiscriminately used, adjuvant chemotherapy increases survival in only a small portion of patients with stage II CRC, but is unnecessary in up to 70% of patients who might be cured by surgery alone. According to a recent Cochrane systematic review, adjuvant therapy improves overall survival (OS) in <10% of patients with stage II colon cancer. However, retrospective analyses have found that certain high-risk features in stage II tumors, such as T4 lesion, poorly differentiated histology, right-sided tumor, lymphovascular invasion (LVI), and tumor perforation, might correlate with greater benefit from adjuvant treatment. Therefore, a salient issue in CRC treatment is how to differentiate patients with high-risk stage II CRC who will benefit from adjuvant chemotherapy from patients with low-risk stage II CRC who will unnecessarily experience adverse effects associated with chemotherapy.

Substantial research has been performed to identify predictors of treatment responses and survival in patients with CRC, and a great deal of recent attention has focused on tumor-based genetic and epigenetic biomarkers. Gao et al analyzed 13 retrospective studies and found that 8 cancer hallmark-based gene signatures could help identify patients with stage II CRC who may benefit from adjuvant chemotherapy. A recent study also showed that a 6-microRNA–based classifier predicted recurrence in patients with stage II CRC, although a model including this classifier and routine clinical prognostic factors only provided moderate discrimination. Despite their potential prognostic values in clinics, the applicability of these models is limited by several caveats, such as the commonly observed intratumoral heterogeneity and the prohibitive costs and specialized equipment and expertise associated with conducting molecular/genomic assays. Thus, it is of significant clinical value to develop novel predictive and prognostic models that are based on simple, low-cost, stable markers.

Several routine blood test–based laboratory variables have been evaluated as prognostic factors in CRC; however, no consensus has yet been reached on their applicability in clinics. Among these variables, carcinoembryonic antigen (CEA) has been used widely as a marker of CRC recurrence, but with suboptimal sensitivity and specificity. Furthermore, no study has systemically analyzed a comprehensive panel of laboratory variables. Given the fact that the role of routine laboratory tests in predicting prognosis has been largely overlooked in clinical settings, we conducted an exploratory analysis using a large patient cohort from a single institution to develop a prognostic model incorporating a comprehensive panel of routine blood test–based laboratory variables, and to explore its prognostic and predictive values in patients with stage II CRC.

Methods

Study Population

Subjects were identified from a hospital-based cohort of patients with histologically confirmed CRC who were treated at Thomas Jefferson University Hospital with an initial diagnosis date from 1998 to 2010. From a cohort of 5,106 patients with CRC, we restricted our study population to those with stage II adenocarcinoma who were surgically treated and had at least one routine blood test within 1 month before surgical removal of their primary tumor. Those patients who were followed for <1 month or without information of recurrence were excluded. This study was approved by the Thomas Jefferson University Hospital Institutional Review Board.

Data Collection

Demographic (age, sex, ethnicity, smoking status, drinking status, and body mass index [BMI]) and basic clinical information (T stage, tumor grade, primary tumor site, LVI, bowel perforation/obstruction, treatments, and recurrence status) were obtained from tumor registry and medical charts (supplemental eTable 1, available with this article at JNCCN.org). Routine blood-based laboratory variables, including tumor marker panel, CBC panel, and comprehensive metabolic panel, which were shown to have prognostic value in previous studies, were considered as candidate variables. If several laboratory variables in a panel were highly correlated (e.g., hemoglobin [Hb]/RBC count/hematocrit, creatinine/urea nitrogen), only the one most clinically relevant and of practical value was included. Candidate laboratory variables were obtained by reviewing medical charts, and included CEA, Hb, albumin, total bilirubin, alanine aminotransferase, aspartate aminotransferase, alka-
line phosphatase, and creatinine levels and red cell distribution width, mean corpuscular volume, WBC count, and platelet count (supplemental eTable 2). The mean value of the laboratory variables was used if multiple measurements were available within 1 month before surgery.

**Statistical Analysis**

SAS 9.4 (SAS Institute Inc.) and STATA, version 12 (StataCorp LP) software packages were used. All statistical tests were 2-sided, and a P value of <0.05 was considered statistically significant. The clinical end point analyzed in this study was disease-free survival (DFS), which was defined as time from surgery to the first event of either recurrence or death from any cause. Patients who were alive and recurrence-free at the last follow-up were censored.

Two-thirds of the patients were randomly selected as a training set to establish the prognostic model, and the remaining patients formed a testing set to assess the predictive performance of the model. Univariate Cox proportional hazards regression and Kaplan-Meier analyses were used to evaluate the association between each variable and patient DFS. Significant demographic and clinical variables in the univariate analyses (P<.05) were forced into the multivariate model; that is, these variables were entered into the initial model and not subject to deletion from the model. A laboratory variable was included in the multivariate stepwise selection only if it demonstrated associations with DFS (P<.05) in all the univariate tests, including continuous, categorical, and log-rank analyses. Continuous laboratory variables with a skewed distribution were transformed toward normality. Moreover, all laboratory variables were kept continuous in the multivariate model construction to avoid loss of power and residual confounding. A laboratory variable was included in the multivariate stepwise selection only if it demonstrated associations with DFS (P<.05) in all the univariate tests, including continuous, categorical, and log-rank analyses. Continuous laboratory variables with a skewed distribution were transformed toward normality. Moreover, all laboratory variables were kept continuous in the multivariate model construction to avoid loss of power and residual confounding. Before model construction, missing data of laboratory variables in the training set were imputed using a multiple imputation method; 10 imputations of the training set were generated based on multivariate normal imputation. The laboratory variables that were significant in univariate analyses were then subject to selection and deletion by a forward stepwise selection using the Akaike information criterion (AIC), and the model with the smallest AIC was selected as the best model for each imputed data set. The final model was composed of the significant demographic and clinical variables, as well as the laboratory variables selected in at least 9 imputed data sets. The parameter estimation was calculated based on the pooled imputed data sets. A prognostic score was then developed for each patient by calculating the sum of each variable weighted by corresponding regression coefficient in the final Cox model.

Discriminatory accuracy for predicting DFS after CRC resection was assessed by constructing receiver operating characteristic (ROC) curves and calculating the area under the curve (AUC). The prediction power was also estimated using Concordance-index (C-index), which works as an extension of AUC to the case of censored survival data. Patients with CRC were divided into low-, medium-, and high-risk groups according to tertile distribution of the prognostic scores. Survival curves for the 3 risk groups were plotted using the Kaplan-Meier method and compared using the log-rank test. The association between prognostic score and patient DFS was estimated using the Cox proportional hazards model. The performance of the final model was tested in the testing set.

**Results**

**Patient Characteristics**

A total of 422 patients with stage II CRC who met the selection criteria were randomly divided into a training set (n=281) and a testing set (n=141). Median follow-up time was 59.64 and 55.81 months, respectively (P=.85). The differences between the training and testing sets with respect to demographic, basic clinical, and laboratory variables were not statistically significant (supplemental eTables 1 and 2).

**Univariate Analysis**

Among the demographic and clinical variables in the training set, age, T stage, and status of LVI and bowel perforation/obstruction were associated with DFS (Table 1). The risk of recurrence after resection doubled in patients aged >67 years (HR, 2.03; P=.0003). Significantly increased recurrence risk was also observed in patients with a T4 stage tumor, LVI, or bowel perforation/obstruction (HR, 1.86, 2.39, and 1.69, respectively). Although black patients, current smokers, obese patients, and patients with right-sided colon cancers had a slightly
increased risk of recurrence, and treatments seemed to prolong survival for these patients, the differences were not statistically significant (supplemental eTable 3). Univariate analysis results of laboratory variables are shown in supplemental eTable 4. Five laboratory variables (CEA, WBC, Hb, albumin, and creatinine) exhibited significant associations with patient DFS when analyzed as both continuous and categorical variables (Table 1).

Stepwise Selection and Final Model Construction

The significant laboratory variables identified from the univariate analyses were considered the candidate predictors for the stepwise selection. In order to address the issue of missing values in laboratory variables, 10 imputed data sets were generated using the multiple imputation method. Supplemental eTable 5 shows the number of times that each of these laboratory variables was selected by stepwise selection; 3 laboratory variables (CEA, Hb, creatinine) were selected in at least 9 imputed data sets. Therefore, the final model included 4 significant demographic and clinical variables and 3 selected laboratory variables. Table 2 summarizes the parameters estimated based on the pooled imputed data sets. A prognostic score was then calculated for each patient based on this model.

Prognostic Performance and Model Validation

The performance of the final model to predict DFS status was evaluated by AUCs from ROC curves. Figure 1 exhibits an AUC of 0.75 (95% CI, 0.69–0.81) for the training set and 0.79 (95% CI, 0.72–0.87) for the testing set, indicating a relatively high accuracy in predicting DFS.
model was also assessed by the C-index, which can be interpreted similarly to the AUC. The C-index for predicting 3-year DFS was 0.73 (95% CI, 0.64–0.83) for the training set and 0.75 (95% CI, 0.65–0.85) for the testing set (data not shown). Compared with the model without the laboratory variables, the predictive accuracy of the final model, including 3 selected laboratory variables, was significantly improved (AUC from 0.67 to 0.75; \( P = .005 \) in the training set, Figure 1A; AUC from 0.69 to 0.79; \( P = .012 \) in the testing set, Figure 1B). Moreover, the discrimination power of the final model seemed to be more evident in patients who did not receive chemotherapy than in those who did (Figure 1C, D).

Based on the tertile distribution of the prognostic scores, we classified patients with CRC into low-, medium-, and high-risk groups. The median DFS was significantly different among the 3 risk groups in the training set (\( P < .0001 \); Figure 2A). Compared with patients in the low-risk group, those in the medium- and high-risk group had a shorter median DFS, with an HR of 1.99 (95% CI, 1.16–3.43) and 4.78 (95% CI, 2.87–7.94), respectively (Table 3). Because the risk score was calculated based on the final model already including the 4 demographic and clinical variables, we did not conduct multivariate analyses to adjust these variables. Results derived from the training set were validated in the testing set (Figure 2B and Table 3).

**Utility of the Model to Identify Appropriate Patients for Chemotherapy**

To identify patients with stage II CRC who might benefit from chemotherapy, we further conducted
stratified analyses according to use of chemotherapy. Because of consistent findings in the training and testing sets and the reduced sample size in stratified analyses, we combined training and testing sets to conduct this analysis. As shown in Figure 3, in the low- and medium-risk groups, survival was shorter among the patients who received chemotherapy compared with those who did not (median DFS, 35.05 months; \( P = .007 \); Figure 3C).

**Discussion**

Statistical models based on routine laboratory data have exhibited promising predictive power for the risk of sporadic CRC. However, the prognostic values of such models have not been comprehensively evaluated, especially in patients with stage II CRC, for whom a consensus on whether to use adjuvant chemotherapy has not been reached. Our current study evaluated the associations between patient DFS and laboratory variables from routine and readily available blood tests. Results showed that a model incorporating 3 laboratory variables (CEA, Hb, creatinine) and 4 demographic and clinical variables (age, T stage, LVI, bowel perforation/obstruction) could potentially be used to stratify patients with stage II CRC into different recurrence-risk groups.

As expected, univariate analyses of clinical characteristics showed that T4 stage, LVI, and bowel perforation/obstruction were associated with a poor DFS in patients with stage II CRC. However, we did not identify a significant association between primary tumor site and DFS, although compared with left-sided colon cancers, right-sided colon cancers and rectosigmoid junction/rectum cancers exhibited a nonsignificantly increased risk of recurrence (supplemental eTable 3). Previous studies showed that patients with metastatic CRC and left-sided tumors had markedly superior progression-free survival, OS, and response rates than those with right-
Figure 3. Kaplan-Meier curves of disease-free survival (DFS) between patients with colorectal cancer who received chemotherapy and those who did not in the (A) low-risk, (B) medium-risk, and (C) high-risk groups. Abbreviations: chemo, chemotherapy; NR, not reached.

sided tumors, but the data for patients with non-metastatic CRC were not consistent. That is, some studies failed to identify tumor site as an independent predictor for DFS, whereas other studies reported worse survival rates associated with left-sided versus right-sided CRC. Moreover, we did not find that BMI predicted recurrence in patients with stage II disease, although obese patients had a non-significantly increased risk for recurrence. These results might also be due to the moderate sample size in our current study, and the further-reduced sample size in the stratified analyses.

The links between the 3 laboratory variables identified in our study and patient DFS are physiologically plausible. Serum CEA, the most frequently used and well-recognized prognostic marker for CRC, has been used for several decades in clinics, although with suboptimal performance. A recent meta-analysis showed an AUC of 0.75 for using CEA levels to predict recurrence of all stages of CRC. In our study, the AUC was 0.62 in the CEA-only model (data not shown), similar to previous reports; however, the performance increased when combined with other significant variables, suggesting the importance of combining multiple routine laboratory variables. Cancer-related anemia is due to multiple etiologies, including chemotherapy-induced myelosuppression, blood loss, functional iron deficiency, erythropoietin deficiency due to renal disease, marrow involvement with tumor, and chronic diseases.

Previous studies showed that anemia was associated with prognosis of cancers, including CRC. Preoperatively, patients with CRC often have anemia because of bleeding from their tumors. It was not surprising that in our study, a high level of Hb was a protective factor against recurrence, and that this variable was selected in the final prediction model. Our model also indicated that patients with a high creatinine value had an unfavorable prognosis. Serum creatinine level is an important indicator of renal function. Reduced creatinine clearance and increased creatinine level were often observed when patients with advanced CRC had toxicities from a chemotherapy agent such as capecitabine. However, lack of information on treatment sequence and chemotherapy agents in our study prevents us from further discussing the possible reason for this association in patients with stage II CRC.
Recent elegant studies showed that models based on molecular and genomic markers had the potential for predicting CRC survivals, but the performance of these markers was moderate for patients with stage II and III CRC.\textsuperscript{10,38} Several genomic predictors in CRC prognosis had AUCs ranging from 0.47 to 0.69.\textsuperscript{39,40} It has thus been suggested that the predictive performance of a model may be significantly improved using an integrative strategy that combines data from multiple dimensions.\textsuperscript{41} Nonetheless, such an integrative strategy based on molecular and genomic assays significantly increases the costs and complexity and reduces the clinical applicability of the model. Our model, which mainly relies on routine laboratory variables, has no incremental cost but still exhibits promising performance. Nonetheless, it is important to note that although a laboratory-based prognostic model could provide similar performance, it should not be equated or compared with molecular-based models. This is because cancer is increasingly recognized as being molecularly driven, and thus, with the reduced costs and increased efficiency of molecular assays, we will increasingly witness cancers being stratified and prognosticated by molecular profiles in the near future.

One important issue in CRC treatment is how to differentiate patients with high-risk stage II CRC who are more likely to benefit from chemotherapy from those who are unlikely to experience significant benefits. In our analysis, the patients in the high-risk group significantly benefited from chemotherapy, whereas chemotherapy seemed to confer an unfavorable effect to those in the low- and medium-risk groups. Similar findings were recently reported by Gao et al,\textsuperscript{8} who used a gene signature–based prognostic score to stratify patients into different risk groups. However, our prediction model is simpler and less expensive, while still providing similar predictive performance in identifying patients with high-risk stage II CRC who are more likely to benefit from chemotherapy use, and is thus more clinically feasible until further advances in molecular stratification are made for prognosis to make it more feasible and affordable. However, it should be noted that although we identified a beneficial effect of chemotherapy in the high-risk group, our model could not discriminate between high- and low-risk patients who received chemotherapy with respect to DFS. Therefore, it should be emphasized that our model is prognostic, but not predictive of benefits from chemotherapy.

The strengths of our study include a highly homogeneous population of patients with stage II CRC, use of training and testing sets to help control model overfitting, use of readily available and clinically relevant laboratory values, and the potential application to identify high-risk patients for chemotherapy treatment. Moreover, the blood tests routinely used in clinics are more reliable and stable than most assays conducted in a biological laboratory, and no specialized equipment or expertise is needed.\textsuperscript{20}

Our study has several limitations. First, we have a moderate sample size, and both the training and testing sets were from a single institution. Further external validations using independent cohorts with a larger sample size are required to test our findings and investigate whether the current model could be extended to patients with CRC at a different tumor stage or patients with other cancers or at another institution. Second, the regimens including chemotherapy have been modified in a large timespan of >10 years; however, the specific agents were not recorded in our cancer registry data. Although we conducted a subcohort analysis and observed more evident benefit in the high-risk patients who were diagnosed in or after 2004 (oxaliplatin was approved in 2004) (supplemental eFigure 1), further investigations with a more precise definition of “risk” are warranted to identify specific groups of patients who will really benefit from chemotherapy. Moreover, we combined training and testing sets together to explore the effect of chemotherapy on recurrence in different risk groups. Although this method increased sample size and statistical power, the observed chemotherapy benefit in the high-risk group is preliminary and exploratory, and should be validated independently. Third, some laboratory variables had a missing rate of >20%. The statistical estimations using multiple imputation cannot unbiasedly represent the real distributions of these variables. Moreover, important clinical/pathologic information, such as microsatellite instability status, is missing as well. Therefore, it remains to be determined whether the model performance can be further improved by including additional clinical and laboratory variables with more complete data collection. Last, although analyzing average levels of laboratory variables could reduce the variations from single measurements,\textsuperscript{30} it may not be the optimal method for each laboratory variable. We pre-
viously reported that maximum value of platelet count appeared to have a better performance than average value in predicting OS of patients with surgically resected CRC, which might be why the average value of platelet was not selected in the current study. We conducted a sensitivity analysis using the highest CEA value and the lowest Hb value; CEA and Hb were still retained in the final model, resulting in a slight increase in model performance (data not shown).

Conclusions
Our exploratory analyses suggest that inexpensive and easily accessible laboratory variables from routine blood tests may reveal important prognostic information for patients with stage II CRC. These results are exploratory and require further validation before clinical application. If validated, our model may provide a cost-effective means to help physicians predict patient survival and make decisions about adjuvant treatment.

References
1192


See JNCCN.org for supplemental online content.