Journal of Current Science and Technology, July - September 2025 Vol. 15 No. 3, Article 123

ISSN 2630-0656 (Online)

Cite this article: Masae, M., Nasawat, A., Sangsakul, N., Mansoh, N., & Ruanglertboon, W. (2025). Exploring the clinical characteristics and survival outcomes in colorectal cancer patients in Hatyai hospital: A retrospective cohort study. *Journal of Current Science and Technology*, *15*(3), Article 123. https://doi.org/10.59796/jcst.V15N3.2025.123

Journal of Current Science and Technology

Journal homepage: https://jcst.rsu.ac.th

Exploring the Clinical Characteristics and Survival Outcomes in Colorectal Cancer Patients in Hatyai Hospital: A Retrospective Cohort Study

Makusee Masae[1,2], Ajira Nasawat[3], Natthawut Sangsakul[2], Nutsara Mansoh[2], and Warit Ruanglertboon[1,4,*]

[1]Research Center in Mathematics and Statistics with Applications, Faculty of Science, Prince of Songkla University, Songkhla 90110, Thailand
[2]Statistics Program, Division of Computational Science, Faculty of Science, Prince of Songkla University, Songkhla 90110, Thailand
[3]Department of Internal Medicine, Hatyai Hospital, Songkhla, 90110, Thailand
[4]Pharmacology Program, Division of Health and Applied Sciences, Faculty of Science, Prince of Songkla University, Songkhla, 90110, Thailand

[*]Corresponding author; E-mail: warit.r@psu.ac.th

Received 13 March 2025; Revised 25 April 2025; Accepted 28 April 2025; Published online 15 June 2025

Abstract

Colorectal cancer (CRC) is considered a significant public health concern worldwide, with substantial morbidity and mortality rates. In Thailand, several campaigns have been implemented to address this issue, such as the establishment of local treatment centers. The Cancer Center of Hatyai Hospital (CCHH) is the latest cancer center affiliated with a tertiary public hospital in the southernmost part of Thailand. However, a systematic assessment of cancer treatment outcomes, including those for CRC patients, has yet to be conducted. Therefore, the current study utilized a retrospective analysis approach to elucidate the survival probability of CRC patients treated at CCHH. A secondary data analysis was conducted using electronic medical records (EMRs), and the selected data were validated and filtered by a certified oncologist and pharmacist. Time-to-event analysis was used to model survival probability across subgroups, and visualized using Kaplan-Meier (KM) plots. Additionally, restricted mean survival time (RMST) analysis was performed to estimate the 3-year survival time of this patient cohort, with an estimated survival time of 24.8 months. The univariate Cox proportional hazards (PH) model was used as an exploratory analysis to identify the influence of clinical variables on survival outcomes. Subsequently, a multivariable Cox PH model was constructed with a set of selected variables. T2 tumor status, the presence of distant metastasis, ECOG score of 4, and poorly differentiated tumor were identified as the strongest predictors of reduced survival among the included variables. As such, this study provides practical insights based on real-world data regarding cancer survivorship and the survival outcomes of CRC patients treated at a public hospital. Additionally, it offers a snapshot of the recent implementation of an early diagnosis campaign.

Keywords: *clinical characteristic; colorectal cancer; early diagnosis; electronic medical records; survival analysis*

Abbreviation

Table A1 Abbreviations used throughout the study

Abbreviation	Explanation
ALB	Albumin
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
BMI	Body Mass Index
BUN	Blood Urea Nitrogen
CCHH	Cancer Center of Hatyai Hospital
CEA	Carcinoembryonic Antigen

Table A1 Cont.

Abbreviation	Explanation
CI	Confidence Interval
CR	Creatinine
CRC	Colorectal Cancer
ECOG	Eastern Cooperative Oncology Group
EMRs	Electronic Medical Records
HB	Hemoglobin
HCT	Hematocrit
HR	Hazard Ratio
KM	Kaplan-Meier
LYMP.L	Lymphocytes
M	Metastasis
N	Nodes
OS	Overall Survival
PH	Proportional Hazards
PMN.N	Polymorphonuclear Neutrophils
RMST	Restricted Mean Survival Time
T	Tumor
WBC	White Blood Cells

1. Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer worldwide, following lung and breast cancer, and the second leading cause of cancer-related deaths, with 916,000 cases reported in 2020 (Olyani et al., 2023; Tan et al., 2024). The prevalence and mortality rates of CRC are rising in many low- and middle-income countries, including Thailand, where cases are frequently diagnosed at an advanced stage (Tiankanon et al., 2021; Muhamad et al., 2023; Wongseree et al., 2023). A report by Lohsiriwat et al., (2020) summarized the current status of CRC in Thailand, indicating that CRC ranks as the third most common cancer, accounting for 11% of the country's total cancer burden. Accurately predicting the prognosis of CRC patients enable the development of personalized treatment plans and supports the implementation of effective public health programs. A precise treatment plan can reduce financial burdens, minimize drug side effects, improve patients' quality of life, and increase survival rates (Xie et al., 2024).

Theoretically, the survival rates of cancer patients serve as crucial indicators of the effectiveness of cancer-specific treatments, and the impact of preventive and survivorship programs (Li et al., 2019; Le et al., 2021). In Asian countries, the overall survival (OS) rate for patients with CRC has shown little improvement in recent decades, with the 5-year OS rate remaining around 60% (Moghimi-Dehkordi, & Safaee, 2012; Le et al., 2021). However, survival rates vary quite significantly across countries, ranging from 36.87% in Thailand (Phimha et al., 2019; Le et al., 2021) to 73.00% in Japan (Tamakoshi et al., 2017; Le et al., 2021). Literature indicates that limitations in cancer epidemiology and survival data often stem from the lack of comprehensive surveillance systems and the poor quality of cancer statistics across responsible organizations (Le et al., 2021). Survival analysis, with its core objective of estimating survival, is primarily used to evaluate outcomes and prognosis in oncology studies, including those focused on CRC. It commonly involves graphical presentations, such as Kaplan-Meier (KM) plots, along with statistical inferences to differentiate research outcomes, identify key covariates, and compare survival probabilities across groups. Therefore, enhancing data quality could improve the predictive capabilities of survival analysis and enable further integration with machine learning for more advanced prognostic modeling (Modi et al., 2020; Ali-Gombe et al., 2021).

In Thailand, multiple oncological studies on CRC have provided valuable insights into prognostic factors, particularly in advanced-stage CRC (Laohavinij et al., 2010) and have assessed overall and stage-specific survival among CRC patients (Kittrongsiri et al., 2020). Similarly, other research has attempted to explore the clinical characteristics and disease outcomes of patients with stage I to III CRC to validate factors influencing treatment outcomes and potential complications among patients treated at Siriraj Hospital, Thailand (Techawathanawanna et al., 2012). Several other studies with a similar approach, aiming to achieve a deeper understanding and enhance predictive or prognostic capabilities, have been published focusing on the Thai population, as mentioned in previous literature (Pongnasuwan, & Chantharakhit, 2023). However, studies specifically examining the southern Thai population are still limited.

Hatyai is one of the most populated districts in southern Thailand. As a tertiary hospital with strong ties to a medical school, Hatyai hospital has become a leading medical institution, treating a wide range of patients from across the southern region, including cancer patients. In late 2021, the Cancer Center of Hatyai Hospital (CCHH) was established as a dedicated facility focused on providing treatment for both local and regional cancer patients. As the center is still in its early stages, treatment outcomes for cancer patients, particularly those with CRC, remain unclear. Therefore, this study aims to assess the survival probability of CRC patients treated at CCHH, serving as a preliminary feasibility study. The findings will offer a brief overview of the recently

implemented early diagnosis campaign and the timely initiation of systemic treatment for CRC patients. Additionally, these data could offer a practical model for countries or clinical settings where early intervention is possible but advanced targeted therapies are not widely accessible.

2. Objectives

This study aims to:

1. Estimate the overall survival rate of CRC patients,
2. Examine the clinical characteristics of patients diagnosed with CRC,
3. Identify factors influencing survival outcomes in CRC patients.

3. Materials and Methods

3.1 Data Collection

This retrospective cohort study was conducted at CCHH. Electronic medical records (EMRs) of patients diagnosed with CRC were reviewed and included in the analysis. Data were collected in 2021, tracking patient information from the start of treatment until the completion of the first follow-up after a successful treatment cycle. The EMRs were reviewed, compiled, and validated for completeness by a certified oncologist and pharmacist, and cross-checking by research assistants. The inclusion criteria were: 1) patients diagnosed with CRC; 2) patients enrolled in CCHH; 3) patients with a complete EMRs from the first follow-up visit onward; and 4) patients for whom the date of mortality was clearly recorded in the EMRs, if applicable. The exclusion criteria were: 1) patients with incomplete or missing key clinical data; 2) patients lost to follow-up before any outcome could be observed; and 3) patients whose date of diagnosis or date of event was ambiguous or inconsistently recorded in the EMRs.

3.2 Statistical Analysis

Survival analysis is a key statistical method used to examine the time until a significant event, such as death or disease progression. Two commonly used techniques in survival analysis are the KM method and the Cox proportional hazards (PH) model (Pawar et al., 2022). The KM method estimates the survival probability of CRC patients over time, providing an understanding of how long patients survive at different time points. The survival function is calculated as follows:

$$S(t_i) = S(t_{i-1})(1-d_i/n_i)$$

where:

$S(t_i)$ is the probability of surviving until time t_i, with $S(0)=1$ (the survival probability at the start of the study is 1),

n_i is the number of patients who are alive or cencored just before time t_{i-1},

d_i is the number of events at time t_{i-1}, with $d_0=0$ (the number of events at the start of the study is 0).

By applying this method to CRC patient data, survival probabilities can be determined at various stages of the disease and compared survival across different patient groups, such as patients at early versus advanced stages of CRC. In addition, the Cox PH model can be used to analyze the impact of variables such as age, gender, and biopsy differentiation on survival, to help identify key risk factors influencing patient outcomes. Together, these methods allow for a comprehensive understanding of CRC survival, and provide valuable insights for clinical decision-making and public health strategies.

The Cox PH model equation for CRC patients would be:

$$h(t|X) = h_0(t)exp(\beta_1X_1+\beta_2X_2+...+\beta_nX_n)$$

where:

$h(t|X)$ is the hazard function for CRC patients at time t,

$h_0(t)$ is the baseline hazard (hazard when all predictors are zero),

$\beta_1,\beta_2,...,\beta_n$ are the coefficients for the predictor variables $X_1,X_2,...,X_n$ (e.g., body mass index (BMI), hemoglobin (HB), hematocrit (HCT), white blood cells (WBC), and cancer stage).

The PH assumption of the Cox model was evaluated using the Schoenfeld residual test to determine whether the effects of covariates on hazard rates remained constant over time. Univariate Cox analyses were first conducted, followed by adjusted multivariable analysis to estimate hazard ratios (HRs) while accounting for potential confounders. Variables for the adjusted model were selected based on prior literature, biological plausibility, clinical relevance, and data availability within the study dataset. Although some covariates did not meet the PH assumption or showed no statistical significance in univariate analysis, they were retained in the multivariable model due to clinical importance. Additionally, variables that produced non-interpretable estimates in univariate analysis - such as those affected by sparse data or quasi-separation - were excluded from

univariate reporting but included in the adjusted model when deemed clinically relevant.

Given the 3-year data collection period and substantial patient censoring, this study employed the restricted mean survival time (RMST) approach to accurately capture the 3-year survival outcome. In studies with heavy censoring, especially toward the end of follow-up, median survival time can be difficult to estimate or imprecise. In contrast, RMST provides a more informative measure of average survival within a fixed period.

Descriptive statistics were used to summarize the baseline clinical characteristics of the patients. KM analysis was employed to estimate and plot survival probabilities based on relevant variables. For subgroup analyses, survival distributions were compared using the log-rank test. Exploratory analyses were conducted using univariate Cox PH models to examine the impact of various covariates on survival. A multivariable Cox PH model was subsequently built based on the meaningful clinical variables identified in the prior exploratory analysis and log-rank test, allowing for adjustment of potential confounders. Data processing and analysis were performed using R version 4.4.3 along with the relevant computational packages.

3.3 Ethics Consideration

This research received ethical approval from the Research Ethics Committee of Hatyai Hospital, Songkhla Province, Thailand, ensuring compliance with ethical guidelines and standards (HYH EC 017-68-01).

4. Results

Table 1 presents the baseline characteristics of a cohort of 99 patients, comprising 40 females and 59 males. The mean age of patients was 66.2 years (SD = 12.2), indicating a predominantly older patient population. The median age was 68.0 years, with an age range spanning from 37.0 to 93.0 years. The majority of patients are aged 65 or older (57.6%), while 30.3% fall within the 51-64 age group. The mean BMI was 21.5 kg/m² (SD = 3.8), with a median of 21.5 kg/m² and a range from 14.1 to 33.9 kg/m². Most patients had a normal BMI (54.5%). The mean HB level was 11.3 g/dL, with 26.3% of patients classified as having low HB. HCT averaged 34.1%, with 36.4% below the normal range. The mean ALB level was 3.7 g/dL; 24.2% had low ALB, and 7.1% were unspecified. AST had a mean of 23.2 U/L, with 18.2% showing elevated levels. WBC count averaged 8,580 cells/μL, with 21.2% above the normal range. Mean PMN.N was 66.0%, with 43.4% below normal. LYMP.L averaged 24.1%, with 12.1% classified as low. Regarding lifestyle factors, most patients reported no history of smoking (61.6%) or alcohol consumption (71.7%). A smaller proportion were identified as former smokers (23.2%) or former alcohol users (20.2%), while current smoking and drinking were reported by 11.1% and 6.1% of patients, respectively. These findings suggest a relatively low prevalence of current tobacco and alcohol use in the study population. A significant majority of patients do not have a family history of cancer (93.9%). In terms of cancer staging, 46.5% were at stage III, and 26.3% are at stage IV. The ECOG performance status shows that most patients have a score of 0 or 1, indicating good health status. Biopsy results show that the majority of patients have well differentiated (52.5%) or moderately differentiated (41.4%) adenocarcinomas, while a smaller proportion (4.1%) have poorly differentiated tumors.

Table 1 Baseline demographic and clinical characteristics of colorectal cancer patients treated at CCHH (N = 99)

Characteristics	Overall (N = 99)
Sex	
Female	40 (40.4%)
Male	59 (59.6%)
Age (year)	
Mean (SD)	66.2 (12.2)
Median [Min, Max]	68.0 [37.0, 93.0]
Age group	
≤ 50	12 (12.1%)
51-64	30 (30.3%)
≥ 65	57 (57.6%)
BMI (kg/m²)	
Mean (SD)	21.5 (3.8)
Median [Min, Max]	21.5 [14.1, 33.9]
BMI group	
Normal	54 (54.5%)
Overweight to obese	18 (18.2%)
Underweight	27 (27.3%)
HB (g/dL)	
Mean (SD)	11.3 (2.2)
Median [Min, Max]	11.3 [5.8, 17.1]
HB group	
Low	26 (26.3%)
Normal to high	73 (73.7%)
HCT (%)	
Mean (SD)	34.1 (6.1)
Median [Min, Max]	33.6 [19.1, 56.8]
HCT group	
Low	36 (36.4%)
Normal to high	63 (63.6%)

Table 1 Cont.

Characteristics	Overall (N = 99)
ALB (g/dL)	
Mean (SD)	3.7 (0.6)
Median [Min, Max]	3.8 [1.7, 4.6]
ALB group	
Low	24 (24.2%)
Normal to high	68 (68.7%)
Not specified	7 (7.1%)
AST (U/L)	
Mean (SD)	23.2 (11.1)
Median [Min, Max]	20.5 [1.9, 67.0]
AST group	
Low	74 (74.7%)
Normal to high	18 (18.2%)
Not specified	7 (7.1%)
WBC (cells/µL)	
Mean (SD)	8,580 (3,710)
Median [Min, Max]	7,520 [3,330, 25,200]
WBC group	
Low	78 (78.8%)
Normal to high	21 (21.2%)
PMN.N (%)	
Mean (SD)	66.0 (12.9)
Median [Min, Max]	66.0 [29.0, 95.0]
PMN.N group	
Low	43 (43.4%)
Normal to high	55 (55.6%)
Not specified	1 (1.0%)
LYMP.L (%)	
Mean (SD)	24.1 (11.9)
Median [Min, Max]	24.0 [2.0, 54.0]
LYMP.L group	
Low	12 (12.1%)
Normal to high	87 (87.9%)

Table 1 Cont.

Characteristics	Overall (N = 99)
Smoking group	
Current	11 (11.1%)
No	61 (61.6%)
Former	23 (23.2%)
Not specified	4 (4.1%)
Alcohol drinking group	
Current	6 (6.1%)
No	71 (71.7%)
Former	20 (20.2%)
Not specified	2 (2.0%)
Cancer stage	
I	6 (6.1%)
II	21 (21.2%)
III	46 (46.5%)
IV	26 (26.3%)
ECOG score	
0	37 (37.4%)
1	46 (46.5%)
2	8 (8.1%)
4	3 (3.0%)
Not specified	5 (5.0%)
Biopsy adenocarcinoma specify group	
Well differentiated	52 (52.5%)
Moderately differentiated	41 (41.4%)
Poorly differentiated	4 (4.1%)
Not specified	2 (2.0%)
Family history of any cancer	
No	93 (93.9%)
Yes	4 (4.1%)
Not specified	2 (2.0%)

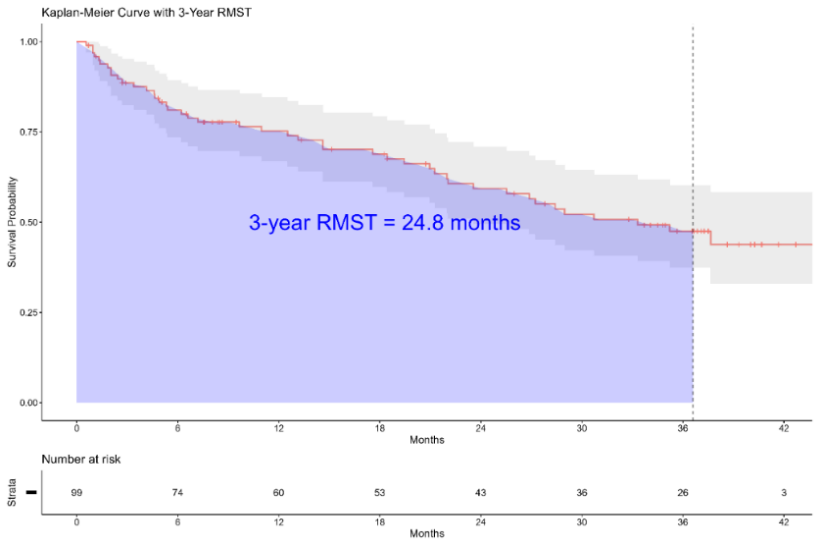

Figure 1 Kaplan-Meier survival curve for the entire colorectal cancer patient cohort (N = 99), with restricted mean survival time (RMST) calculated over a 3-year follow-up. Patient-at-risk table shown beneath the curve.

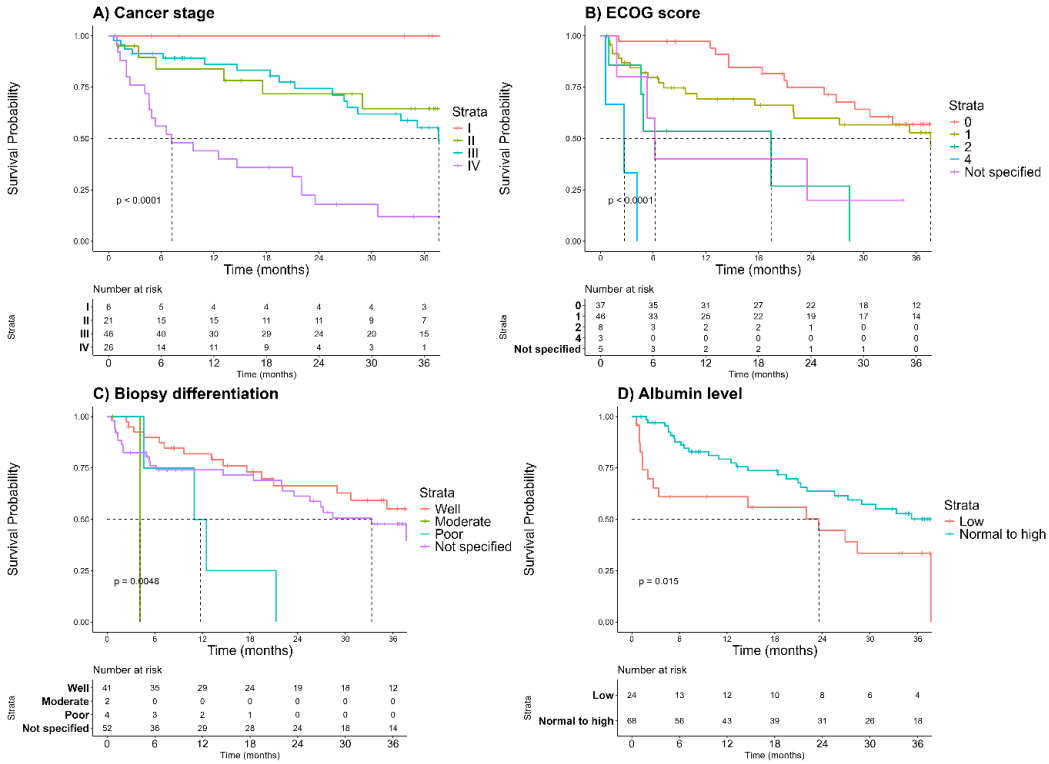

Figure 2 Kaplan-Meier survival curves stratified by significant clinical factors: (A) Cancer stage, (B) ECOG performance score, (C) Tumor biopsy differentiation, and (D) Serum albumin level. P-values derived from log-rank tests.

Figure 1 illustrates the survival probability of the patient cohort using the RMST approach, with the survival curve indicating that more than 50% of patients remained alive over the 3-year follow-up period. According to the analysis, the 3-year RMST for this patient cohort was 24.8 months. The number of patients at risk declined from 99 at baseline to 74 at 6 months, 60 at 12 months, 43 at 24 months, and 3 at 42 months, reflecting events and censoring over time. Subsequent subgroup analyses based on the log-rank test (Figure 2) revealed significant differences in survival probabilities associated with variables such as albumin level, cancer stage at baseline, ECOG performance status, and biopsy differentiation.

As described in the Methods, the PH assumption was evaluated using the Schoenfeld residual test. Supplementary Figure S1 provides an example of the Schoenfeld residual test applied to several variables. However, quasi-separation was observed in the cancer stage variable, primarily due to the absence of events in one subgroup (stage I). This resulted in non-interpretable HR estimates in both univariate and multivariable Cox models. Therefore, the cancer stage was excluded from the regression analyses to avoid misleading conclusions. To address the exclusion of this variable from the regression models, the 3-year RMST was additionally calculated for each cancer stage subgroup, as well as for subgroups of other relevant variables, and the results are presented in Supplementary Table S1. This offers further numerical insight into survival differences across subgroups.

Table 2 summarizes the HRs and adjusted HRs (aHRs) for the variables analyzed in this study. In this analysis, significant associations with increased HR were identified for ECOG scores of 2 (HR = 4.71, 95% CI: 1.64 - 13.5; p = 0.004) and 4 (HR = 23.1, 95% CI: 5.74 - 92.6; p < 0.001), as well as unspecified ECOG scores (HR = 3.84, 95% CI: 1.24 - 11.9; p = 0.019). Additional significant associations were observed for poorly differentiated biopsy results (HR = 4.75, 95% CI: 1.54 - 14.7; p = 0.007), unspecified biopsy differentiation (HR = 9.64, 95% CI: 1.19 - 77.8; p = 0.033), N2 lymph node status (HR = 2.50, 95% CI: 1.12 - 5.55; p = 0.025), presence of metastasis (M1) (HR = 4.76, 95% CI: 2.58 - 8.78; p < 0.001), normal-to-high WBC counts (HR = 2.37, 95% CI: 1.24 - 4.54; p = 0.009), and normal-to-high

PMN.N levels (HR = 1.87, 95% CI: 1.01 - 3.48; p = 0.048). Conversely, normal-to-high ALB levels were significantly associated with a reduced risk of events (HR = 0.46, 95% CI: 0.25 - 0.87; p = 0.017).

In the adjusted analysis, aHRs were estimated by including relevant variables, as described in the Methods section. Factors that remained independently associated with increased hazard were ECOG scores of 4 (aHR = 17.5, 95 % CI: 4.50 - 68.0; p < 0.001), poor biopsy differentiation (aHR = 7.73, 95 % CI: 2.65 - 22.5; p < 0.001), T2 status (aHR = 25.4, 95 % CI: 6.98 - 92.4; p < 0.001) and M1 status (aHR = 17.7, 95 % CI: 9.02 - 34.7; p < 0.001). Additionally, several other factors reached statistical significance only after adjustment. These included male sex (aHR = 0.36, 95% CI: 0.18 - 0.73; p = 0.005), older age groups (51 - 64 years: aHR = 5.89, 95% CI: 2.92 - 11.9; ≥ 65 years: aHR = 12.6, 95% CI: 6.53 - 24.2; both p < 0.001), normal-to-high HCT levels (aHR = 0.39, 95% CI: 0.21 - 0.74; p = 0.004), normal-to-high ALB levels (aHR = 0.25, 95% CI: 0.13 - 0.50; p < 0.001), normal-to-high PMN.N levels (aHR = 6.85, 95% CI: 3.39 - 13.8; p < 0.001), and N2 status (aHR = 2.07, 95% CI: 1.07 - 3.98; p = 0.03).

5. Discussion

Cancer remains one of the most challenging non-communicable diseases that humankind has yet to fully overcome. A major factor hindering treatment outcomes is cancer heterogeneity and its dynamic evolution over time. Cancer types can vary significantly in molecular biology, resulting in diverse treatment responses. Our understanding of each cancer type may not keep pace with the global cancer burden observed across communities. Accordingly, this CRC-focused study adopts a bottom-up approach to assess overall survival probability and explore potential phenotypic markers may predict disease outcomes.

CRC patients enrolled at CCHH, one of the newest cancer centers at the southernmost government-funded institution, demonstrated a 3-year RMST of 24.8 months, with approximately 50% of patients surviving up to three years. Interestingly, the patient cohort had an equal gender distribution, despite statistics showing that men are at higher risk of developing this cancer than women. Nevertheless, the overall survival rates reported in this study align with findings from several previous studies (Monkhan et al., 2023). However, variations in survival rates were observed across different subgroups. The log-rank test revealed significant differences in survival probabilities, particularly associated with ALB level and ECOG score. Furthermore, an exploratory univariate Cox PH analysis demonstrated that various baseline characteristics significantly influenced the risk of events, as detailed in the results section.

Table 2 Univariate and multivariable Cox proportional hazards regression analyses of factors associated with survival in colorectal cancer patients.

Characteristics	Univariate			Multivariable		
	HR	95% CI	p-value	aHR	95% CI	p-value
Sex						
Female	1.00	-		1.00	-	
Male	0.91	0.50, 1.66	0.8	0.36	0.18, 0.73	0.005
Age group						
≤ 50	1.00	-		1.00	-	
51-64	0.88	0.28, 2.72	0.8	5.89	2.92, 11.9	< 0.001
≥ 65	1.29	0.45, 3.68	0.6	12.6	6.53, 24.2	< 0.001
ECOG score						
0	1.00	-		1.00	-	
1	1.50	0.74, 3.04	0.3	0.81	0.43, 1.54	0.5
2	4.71	1.64, 13.5	0.004	1.78	0.67, 4.70	0.2
4	23.1	5.74, 92.6	< 0.001	17.5	4.50, 68.0	< 0.001
Not specified	3.84	1.24, 11.9	0.019	1.18	0.41, 3.41	0.8
Smoking group						
No	1.00	-				
Current	0.90	0.34, 2.32	0.8		Not included	
Former	1.02	0.51, 2.06	> 0.9			
Not specified	0.00	0.00, Inf	> 0.9			

Table 2 Cont.

Characteristics	Univariate			Multivariable		
	HR	95% CI	p-value	aHR	95% CI	p-value
Alcohol drinking group						
No	1.00	-				
Current	0.94	0.29, 3.08	> 0.9		Not included	
Former	1.36	0.66, 2.78	0.4			
Not specified	0.00	0.00, Inf	> 0.9			
Family history of any cancer						
No	1.00	-		1.00	-	
Yes	0.39	0.05, 2.82	0.3	1.7	0.23, 12.9	0.6
Not specified	0.00	0.00, Inf	> 0.9	0.00	0.00, Inf	> 0.9
Biopsy adenocarcinoma specify group						
Moderately differentiated	1.00	-		1.00	-	
Well differentiated	1.43	0.75, 2.73	0.3	2.42	1.28, 4.58	0.007
Poorly differentiated	4.75	1.54, 14.7	0.007	7.73	2.65, 22.5	< 0.001
Not specified	9.64	1.19, 77.8	0.033	0.06	0.01, 0.51	0.01
T						
T1	1.00	-		1.00	-	
T2	3.20	0.33, 31.0	0.3	25.40	6.98, 92.4	< 0.001
T3	2.19	0.30, 16.2	0.4	3.44	1.78, 6.63	< 0.001
T4	4.00	0.53, 30.3	0.2	5.13	2.61, 10.1	< 0.001
N						
N0	1.00	-		1.00	-	
N1	1.74	0.70, 4.35	0.2	0.87	0.41, 1.88	0.7
N2	2.50	1.12, 5.55	0.025	2.07	1.07, 3.98	0.03
M						
M0	1.00	-		1.00	-	
M1	4.76	2.58, 8.78	< 0.001	17.7	9.02, 34.7	< 0.001
BMI group						
Normal	1.00	-		1.00	-	
Overweight to obese	1.88	0.85, 4.13	0.12	1.11	0.52, 2.37	0.8
Underweight	1.77	0.90, 3.49	0.1	0.58	0.29, 1.16	0.12
HB group						
Low	1.00	-		1.00	-	
Normal to high	0.66	0.35, 1.27	0.2	1.74	0.87, 3.50	0.12
HCT group						
Low	1.00	-		1.00	-	
Normal to high	0.56	0.31, 1.02	0.057	0.39	0.21, 0.74	0.004
WBC group						
Low	1.00	-		1.00	-	
Normal to high	2.37	1.24, 4.54	0.009	0.6	0.27, 1.35	0.2
PMN.N group						
Low	1.00	-		1.00	-	
Normal to high	1.87	1.01, 3.48	0.048	6.85	3.39, 13.8	< 0.001
LYMP.L group						
Low	1.00	-		1.00	-	
Normal to high	0.71	0.30, 1.68	0.4	0.68	0.25, 1.81	0.4
Monocyte group						
Low	1.00	-			Not included	
Normal to high	0.60	0.23, 1.52	0.3			
Eosinophil group						
Low	1.00	-			Not included	
Normal to high	0.86	0.46, 1.58	0.6			

Table 2 Cont.

Characteristics	Univariate			Multivariable		
	HR	95% CI	p-value	aHR	95% CI	p-value
BUN group						
Low	1.00	-			Not included	
Normal to high	0.72	0.36, 1.45	0.4			
Creatinine group						
Low	1.00	-			Not included	
Normal to high	0.62	0.30, 1.25	0.2			
ALB group						
Low	1.00	-		1.00	-	
Normal to high	0.46	0.25, 0.87	0.017	0.25	0.13, 0.50	< 0.001
AST group						
Low	1.00	-			Not included	
Normal to high	1.14	0.53, 2.48	0.7			
ALT group						
Low	1.00	-			Not included	
Normal to high	0.38	0.12, 1.23	0.11			
CEA group						
Low	1.00	-			Not included	
Normal to high	1.23	0.56, 2.71	0.6			

Note: HR = hazard ratio, aHR = adjusted hazard ratio, CI = confidence interval

The association between elevated WBC and PMN.N levels aligns with previous literature, as these parameters likely reflect the occurrence of inflammation in cancer-related inflammatory processes (Weng et al., 2022). Elevated WBC has been reported as a poor prognostic factor in several cancer types, often linked to unfavorable outcomes (Mabuchi et al., 2011). On the other hand, the ECOG score is used to assess a patient's ability to carry out daily activities while living with cancer. In this study, higher ECOG scores (greater than 1) indicate increased frailty and a greater likelihood of poorer survival outcomes. This finding serves as concrete evidence-consistent with previous literature in cancer patients-that a low ECOG score is associated with better survival outcomes and functions as an independent prognostic factor in this patient cohort (Haus et al., 2020). However, one particular group-where the ECOG score was not specified exhibited a significant association with worse outcomes in the univariate Cox PH analysis (HR = 3.84, 95% CI: 1.24 - 11.9, p = 0.019), although this association was not significant in the adjusted analysis. Additionally, the log-rank test for this group showed a trend toward poorer survival probability compared to patients with an ECOG score lower than 1. These patients may have corresponded to ECOG 3, though no definitive evidence supports this assumption. The absence of ECOG scores and other unspecified variables reflects the limitations inherent in real-world data, where complete data collection cannot always be guaranteed (Grimberg et al., 2021; Tang et al., 2023). However, this limitation also presents an opportunity to advocate for improved data curation by the organization that owns the data. Enhanced data collection will be crucial for improving predictive accuracy and enabling more robust data utilization in future analyses.

Following the univariate Cox PH analysis, an adjusted multivariable analysis was conducted. Several variables remained significant in the adjusted analysis, including ECOG score of 4, poorly differentiated tumor biopsy, metastasis status, and T2 tumor status, each significantly associated with increased hazard. However, some variables-such as T2 tumor status-produced imprecise aHRs, as indicated by extremely wide 95% CI, likely due to quasi-separation caused by a very small number of events in the reference group. To provide additional insight and address limitations outlined in the Results section, the 3-year RMST was calculated for each variable in the dataset. These findings are presented in Supplementary Table S1. Accordingly, interpretation of the aHRs should be interpreted with caution, particularly in cases where data sparsity may have compromised estimate precision. Notably, some variables that were not statistically significant in univariate analysis became significant in the multivariable model, likely due to the adjustment for confounding effects. This highlights the importance

of accounting for potential interactions among variables, especially in the context of real-world data, where complex relationships may not be apparent in isolated comparisons. Nevertheless, based on the findings from the adjusted analysis, the Cox regression model identified the top four strongest predictors of increased mortality, ranked from highest to lowest aHR, as T2 tumor status, distant metastasis, ECOG score of 4, and poorly differentiated tumor biopsy. In contrast, several variables, such as WBC count and ALB level, that showed significant associations in the univariate analysis did not retain significance in the adjusted model, suggesting that their apparent effects may have been confounded by other factors.

The strength of the current study lies in its use of real-world data, offering findings that accurately reflect the current patient demographic. The southern Thai population exhibits distinct characteristics compared to other regions of the country. Therefore, these findings may support the development of tailored public health strategies for this population, enabling more precise and effective healthcare planning and policy implementation. Long-term data collection with ongoing survival analysis is recommended to further explore additional dimensions of this patient cohort and others treated at the cancer center. Additionally, implementing machine learning once sufficient data are accumulated could yield deeper clinical insights, enhancing decision-making and personalized care.

Despite the aforementioned strengths, it must be acknowledged that this study has several limitations. First, the relatively small sample size may limit the statistical power and generalizability of the findings. This limitation partly stems from the single-center design, which restricted the pool of eligible patients. Second, the three-year data collection period may not capture long-term survival trends or disease progression. A multicenter study involving a larger and more diverse population would be essential to provide a broader perspective and enable adjustments for region-specific differences, thereby strengthening the validity and applicability of future research.

6. Conclusion

CRC patients treated at CCHH exhibited prognostic factors consistent with prior literature, such as ECOG score and baseline WBC. The 3-year RMST for this patient cohort was 24.8 months. This analysis identified several predictors strongly associated with reduced survival, including T2 tumor status, distant metastasis, ECOG score of 4, and poorly differentiated tumor biopsy. Despite these findings, careful consideration is warranted when interpreting certain estimates, as wide confidence intervals reflect limited precision due to sparsity in the reference group.

7. Acknowledgements

The authors would like to express their deepest gratitude to the healthcare professionals at CCHH for their support in facilitating data collection and other related matters.

8. References

Ali-Gombe, M., Mustapha, M. I., Folasire, A., Ntekim, A., & Campbell, O. B. (2021). Pattern of survival of breast cancer patients in a tertiary hospital in South West Nigeria. *Ecancermedicalscience*, *15*, Article 1192. https://doi.org/10.3332/ecancer.2021.1192

Grimberg, F., Asprion, P. M., Schneider, B., Miho, E., Babrak, L., & Habbabeh, A. (2021). The real-world data challenges radar: a review on the challenges and risks regarding the use of real-world data. *Digital Biomarkers*, *5*(2), 148-157. https://doi.org/10.1159/000516178

Haus, R., Janssen, S., Schild, S. E., & Rades, D. (2020). Eastern cooperative oncology group performance score is associated with survival after radiotherapy of bone metastases from prostate cancer. *in vivo*, *34*(2), 679-682. https://doi.org/10.21873/invivo.11823

Kittrongsiri, K., Wanitsuwan, W., Prechawittayakul, P., Sangroongruangsri, S., Cairns, J., & Chaikledkaew, U. (2020). Survival analysis of colorectal cancer patients in a Thai hospital-based cancer registry. *Expert Review of Gastroenterology & Hepatology*, *14*(4), 291-300. https://doi.org/10.1080/17474124.2020.1740087

Laohavinij, S., Maneechavakajorn, J., & Techatanol, P. (2010). Prognostic factors for survival in colorectal cancer patients. *Journal of the Medical Association of Thailand*, *93*(10), 1156–1166.

Le, D. D., Van Vo, T., & Sarakarn, P. (2021). Overall survival rate of Vietnamese patients with colorectal cancer: A hospital-based cohort study in the central region of Vietnam. *Asian Pacific Journal of Cancer Prevention: APJCP*, *22*(11), 3569–3575. https://doi.org/10.31557/APJCP.2021.22.11.3569

Li, X., Zhou, Y., Luo, Z., Gu, Y. A., Chen, Y., Yang, C., ... & Zhao, G. (2019). The impact of screening on the survival of colorectal cancer in Shanghai, China: A population based study. *BMC Public Health*, *19*, 1-9. https://doi.org/10.1186/s12889-019-7318-8

Lohsiriwat, V., Chaisomboon, N., & Pattana-Arun, J. (2020). Current colorectal cancer in Thailand. *Annals of Coloproctology*, *36*(2), 78-82. https://doi.org/10.3393/ac.2020.01.07

Mabuchi, S., Matsumoto, Y., Isohashi, F., Yoshioka, Y., Ohashi, H., Morii, E., ... & Kimura, T. (2011). Pretreatment leukocytosis is an indicator of poor prognosis in patients with cervical cancer. *Gynecologic Oncology*, *122*(1), 25-32. https://doi.org/10.1016/j.ygyno.2011.03.037

Modi, N. D., Sorich, M. J., Rowland, A., Logan, J. M., McKinnon, R. A., Kichenadasse, G., ... & Hopkins, A. M. (2020). A literature review of treatment-specific clinical prediction models in patients with breast cancer. *Critical Reviews in Oncology/Hematology*, *148*, Article 102908. https://doi.org/10.1016/j.critrevonc.2020.102908

Moghimi-Dehkordi, B., & Safaee, A. (2012). An overview of colorectal cancer survival rates and prognosis in Asia. *World Journal of Gastrointestinal Oncology*, *4*(4), 71-75. https://doi.org/10.4251/wjgo.v4.i4.71

Monkhan, N., Phimha, S., Prasit, N., Senahad, N., & Pinsuwan, C. (2023). Colorectal cancer survival in Thailand. *International Journal of Public Health Asia Pacific*, 2(3), 1-8. https://doi.org/10.62992/ijphap.v2i3.36

Muhamad, N. A., Ma'amor, N. H., Rosli, I. A., Leman, F. N., Abdul Mutalip, M. H., Chan, H. K., ... & Abu Hassan, M. R. (2023). Colorectal cancer survival among Malaysia population: data from the Malaysian National Cancer Registry. *Frontiers in Oncology*, *13*, Article 1132417. https://doi.org/10.3389/fonc.2023.1132417

Olyani, S., Ebrahimipour, H., Taraghdari, M. M., Jamali, J., & Peyman, N. (2023). Colorectal Cancer awareness and related factors among adults attending primary Healthcare in North-Eastern of Iran: a cross-sectional study. *Journal of Research in Health Sciences*, *23*(3), Article e00589. https://doi.org/10.34172/jrhs.2023.124

Pawar, A., Chowdhury, O. R., & Salvi, O. (2022). A narrative review of survival analysis in oncology using R. *Cancer Research, Statistics, and Treatment*, *5*(3), 554-561. https://doi.org/10.4103/crst.crst_230_22

Phimha, S., Promthet, S., Suwanrungruang, K., Chindaprasirt, J., Bouphan, P., Santong, C., & Vatanasapt, P. (2019). Health insurance and colorectal cancer survival in Khon Kaen, Thailand. *Asian Pacific Journal of Cancer Prevention*, *20*(6), Article 1797. https://doi.org/10.31557/APJCP.2019.20.6.1797

Pongnasuwan, W., & Chantharakhit, C. (2023). Survival Analysis and prognostic factors for metastatic colorectal cancer patients treated with chemotherapy. *Journal of the Medical Association of Thailand*, *106*(1). 1-7. https://doi.org/10.35755/jmedassocthai.2023.01.13726

Tamakoshi, A., Nakamura, K., Ukawa, S., Okada, E., Hirata, M., Nagai, A., ... & BioBank Japan Cooperative Hospital Group. (2017). Characteristics and prognosis of Japanese colorectal cancer patients: The BioBank Japan Project. *Journal of Epidemiology*, *27*(3), S36-S42. https://doi.org/10.1016/j.je.2016.12.004

Tan, C., Li, Y., Wang, K., Lin, Y., Chen, Y., & Zheng, X. (2024). Causal roles and clinical utility of cardiovascular proteins in colorectal cancer risk: a multi-modal study integrating mendelian randomization, expression profiling, and survival analysis. *BMC Medical Genomics*, *17*(1), 138. https://doi.org/10.1186/s12920-024-01909-4

Tang, M., Pearson, S. A., Simes, R. J., & Chua, B. H. (2023). Harnessing real-world evidence to advance cancer research. *Current Oncology*, *30*(2), 1844-1859. https://doi.org/10.3390/curroncol30020143

Techawathanawanna, S., Nimmannit, A., & Akewanlop, C. (2012). Clinical characteristics and disease outcome of UICC stages I–III colorectal cancer patients at Siriraj Hospital. *Journal of the Medical Association of Thailand*, *95*(2), S189-S198.

Tiankanon, K., Aniwan, S., & Rerknimitr, R. (2021). Current status of colorectal cancer and its public health burden in Thailand. *Clinical Endoscopy*, *54*(4), 499-504. https://doi.org/10.5946/ce.2020.245-IDEN

Weng, M., Zhao, W., Yue, Y., Guo, M., Nan, K., Liao, Q., ... & Miao, C. (2022). High

preoperative white blood cell count determines poor prognosis and is associated with an immunosuppressive microenvironment in colorectal cancer. *Frontiers in Oncology*, *12*, Article 943423. https://doi.org/10.3389/fonc.2022.943423

Wongseree, P., Hasgul, Z., Leerapan, B., Iramaneerat, C., Phisalprapa, P., & Jalali, M. S. (2023). Dynamics of colorectal cancer screening in low and middle-income countries: A modeling analysis from Thailand. *Preventive Medicine*, *175*, Article 107694. https://doi.org/10.1016/j.ypmed.2023.107694

Xie, H., Wei, L., Tang, S., & Gan, J. (2024). Aminotransferase-to-lymphocyte ratio as a valuable prognostic marker for patients with stage I-III colorectal cancer: a retrospective study. *Frontiers in Oncology*, *14*, Article 1446557. https://doi.org/10.3389/fonc.2024.1446557

Supplementary Data

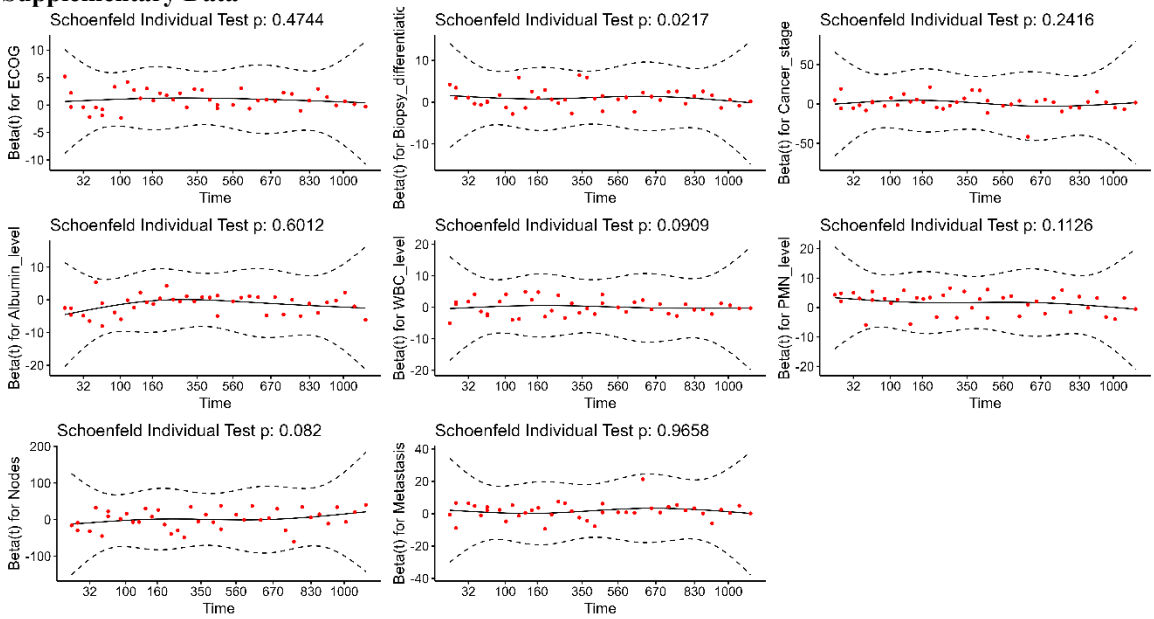

Figure S1 Schoenfeld residual plots used to assess the proportional hazards assumption in selected covariates for Cox regression

Table S1 Subgroup analysis of restricted mean survival time (RMST) over 3 years by clinical and demographic variables

Characteristics	Subgroups	3-year RMST (in months)
Sex	Male	25.6
	Female	22.6
Age group	≤ 50	23.5
	51-64	27.2
	≥ 65	23.1
BMI group	Normal	26.6
	Overweight to obese	21.0
	Underweight	21.4
HB group	Low	20.5
	Normal to high	25.8
HCT group	Low	20.9
	Normal to high	26.5
WBC group	Low	26.9
	Normal to high	15.9
PMN.N group	Low	28.8
	Normal to high	21.5
LYMP.L group	Low	20.0
	Normal to high	25.2
Monocyte group	Low	18.6
	Normal to high	25.1
Eosinophil group	Low	24.2
	Normal to high	25.0
BUN group	Low	22.3
	Normal to high	25.1

Table S1 Cont.

Characteristics	Subgroups	3-year RMST (in months)
Creatinine group	Low	18.5
	Normal to high	25.9
ALB group	Low	19.0
	Normal to high	26.1
AST group	Low	24.4
	Normal to high	22.9
ALT group	Low	23.1
	Normal to high	30.5
CEA group	Low	27.0
	Normal to high	26.1
Cancer stage	I	35.8
	II	27.6
	III	28.2
	IV	13.5
ECOG score	0	29.4
	1	24.5
	2	14.5
	4	2.41
	Not specified	14.3
T	T1	28.7
	T2	19.7
	T3	26.0
	T4	20.7
N	N0	29.2
	N1	23.9
	N2	22.0
M	M0	28.8
	M1	13.7
Smoking group	Current	26.5
	No	23.4
	Former	24.1
	Not specified	35.8
Alcohol drinking group	Current	27.7
	No	24.4
	Former	22.2
	Not specified	35.8
Biopsy adenocarcinoma specify group	Well differentiated	24.2
	Moderately differentiated	27.0
	Poor differentiated	12.3
	Not specified	3.94
Family history of any cancer	Yes	27.1
	No	24.0
	Not specified	35.8