RESEARCH ARTICLE

Editorial Process: Submission:09/11/2024 Acceptance:12/24/2024

Development and Validation of Machine Learning Model Platelet Index-based Predictor for Colorectal Cancer Stage

Citra Aryanti[1]*, Ronald Erasio Lusikooy[2], Samuel Sampetoding[2], Sachraswaty R. Laidding[1], Warsinggih Warsinggih[2], Erwin Syarifuddin[2], Julianus Aboyaman Uwuratuw[2], Muhammad Ihwan Kusuma[1], Ibrahim Labeda[3], Murny Abdul Rauf[4]

Abstract

Introduction: Colorectal cancer (CRC) staging is essential for effective treatment planning and prognosis. While platelet indices have shown promise in indicating CRC aggressiveness, a platelet index-based predictor for CRC staging has not been established in Indonesia. This study aimed to explore the relationship between platelet indices and CRC stage and to develop a predictive model and application. **Methods:** This cross-sectional study analyzed 369 CRC patients from Dr. Wahidin Sudirohusodo Hospital. Key parameters included age, gender, tumor location, and platelet indices: platelet count (PC), mean platelet volume (MPV), platelet distribution width (PDW), plateletcrit, and the MPV/PC ratio. Data were processed using SPSS 25, MATLAB, and Streamlit. **Results and Discussion:** The analysis revealed significant correlations between elevated platelet indices and advanced CRC stages. Various machine learning models were developed, with Support Vector Machine (SVM) achieving the highest accuracy at 82.9%, followed closely by K-Nearest Neighbors (82.7%), Neural Network (81.5%), Naive Bayes (80.5%), and logistic regression (51.5%). The most effective model was implemented as a portable application through Streamlit, yielding 79.2% internal validation and 89.2% external validation. **Conclusion:** This study highlights a significant association between increased platelet indices and advanced CRC stages. The innovative platelet index-based predictor for CRC staging offers promising potential for enhancing individualized clinical decision-making. By providing a non-invasive method that complements existing staging techniques, this approach could significantly improve patient outcomes through earlier and more accurate CRC staging. The findings underscore the importance of integrating simple, accessible biomarkers into clinical practice to enhance diagnostic precision.

Keywords: Prediction- Machine Learning- Platelet Count- Indonesia- Colorectal Neoplasms

Asian Pac J Cancer Prev, **25 (12)**, 4425-4433

Introduction

Colorectal cancer (CRC) remains a significant global health concern, with high morbidity and mortality rates. It ranks as the third most common cancer in men (10.0%) and the second most common in women (9.2%) worldwide [1]. Geographical variations in CRC incidence are notable, with the highest rates observed in Australia and New Zealand, while Central and Eastern Europe report the highest mortality rates [2].

In Indonesia, CRC is the fourth most prevalent cancer, accounting for 8.6% of all cancer cases. According to Globocan 2020 data, there were 396,914 new CRC cases in Indonesia, with 234,411 deaths reported [3]. A study at Wahidin Sudirohusodo Hospital in Makassar revealed that most CRC patients (57.3%) were aged 50 years or older, with a majority presenting at advanced stage IV (34.8%) [4]. The overall 5-year survival rate for CRC in this setting was reported to be 36.5% [5].

The TNM staging system, developed by the American Joint Cancer Committee (AJCC), is widely used to classify CRC progression. It demonstrates a clear correlation between advancing stages and declining 5-year survival rates: 75.6%, 42.1%, 26.5%, and 11.3% for stages I through IV, respectively [6]. The high cancer mortality rate is largely attributed to late detection, often when metastasis has already occurred, significantly limiting effective treatment options. It is estimated that at least 15% of cancer-related deaths within 5 years could be prevented through early detection [7].

The tumor microenvironment plays a crucial role in cancer progression. It comprises a complex matrix

*[1]Departement of Surgery, Faculty of Medicine, Hasanuddin University, Makassar, South Sulawesi, Indonesia. [2]Division of Digestive Surgery, Department of Surgery, Hasanuddin University, Indonesia. [3]Division of Digestive, Department of Surgery, Dr. Wahidin Sudirohusodo General Hospital, Makassar, South Sulawesi, Indonesia. [4]Department of Surgery, Hasanuddin University, Makasar, South Sulawesi, Indonesia. *For Correspondence: citraary91@gmail.com*

of extracellular components, cytokines, growth factors, adhesion molecules, and various cellular elements including fibroblasts, immune cells, adipocytes, pericytes, epithelial cells, lymphatic and endothelial cells, and platelets [8]. The relationship between elevated platelet counts and cancer has been observed since 1872 when Leopold Riess noted increased platelet numbers in patients with malignant diseases [8].

Platelets, small anucleate cytoplasmic fragments derived from megakaryocytes, measure 2-4 μm in size and have a normal count of 150-400 x 10^9/L [9]. Platelet indices, including platelet count (PC), mean platelet volume (MPV), platelet distribution width (PDW), and plateletcrit (PCT), serve as biomarkers of platelet activation and predictors of platelet size, morphology, and proliferation kinetics [10, 11].

The complex bidirectional communication between cancer cells and platelets has been a research subject for over a century. Studies have shown that nearly 40% of patients with solid tumors in the gastrointestinal tract, lungs, breast, ovaries, and prostate exhibit platelet counts exceeding 400,000 mm3 [12].

Given the significant impact of CRC on global health and the potential role of platelets in cancer progression, further investigation into the relationship between platelet indices and CRC staging is warranted. This research explores the association between platelet indices and CRC stages, potentially offering new insights into early detection and prognosis of colorectal cancer.

Materials and Methods

Study Design and Population

This was a cross-sectional study conducted at Dr. Wahidin Sudirohusodo General Hospital in Makassar, Indonesia. The study population included patients diagnosed with colorectal cancer (CRC) who visited the hospital in 2023. Patients with incomplete medical records or those who had received prior treatment were excluded. The institutional ethics committee approved the study protocol.

Sample Size and Sampling Technique

The sample size was calculated using the formula for cross-sectional studies, with a confidence level of 95% and a margin of error of 5%. Based on this calculation, a minimum sample size of 369 patients was determined. Consecutive sampling was used to recruit eligible patients until the required sample size was reached.

Data Collection

Demographic and clinical data were collected from medical records, including age, gender, tumor location, and cancer stage according to the American Joint Committee on Cancer (AJCC) TNM staging system, 8th edition. Blood samples were obtained from all patients prior to any treatment. Platelet indices were measured using an automated hematology analyzer (Sysmex XN-1000, Sysmex Corporation, Kobe, Japan). The following platelet indices were recorded: Platelet count (PC), Mean platelet volume (MPV), Platelet distribution width (PDW), Plateletcrit (PCT), and MPV to PC ratio (MPV/PC).

Statistical Analysis

Data were analyzed using SPSS version 25 (IBM Corp., Armonk, NY, USA). Descriptive statistics were used to summarize patient characteristics. Continuous variables were expressed as mean ± standard deviation or median (interquartile range), depending on the normality of distribution. Categorical variables were presented as frequencies and percentages.

The relationship between platelet indices and the CRC stage was analyzed using one-way ANOVA or Kruskal-Wallis test for continuous variables, depending on data distribution. When appropriate, post-hoc analysis was performed using Tukey's HSD or Dunn's test. Pearson's chi-square test or Fisher's exact test was used for categorical variables.

Receiver operating characteristic (ROC) curve analysis was performed to evaluate the diagnostic performance of platelet indices in predicting advanced CRC stages. The area under the curve (AUC), sensitivity, specificity, and optimal cut-off values were calculated.

Machine Learning Model Development

To develop a predictive model for CRC staging based on platelet indices, various machine-learning algorithms were implemented using MATLAB R2021b (MathWorks, Natick, MA, USA). The following algorithms were evaluated: Logistic Regression, K-Nearest Neighbors (KNN), Naive Bayes, Support Vector Machine (SVM), and Neural Network. The dataset was split into training (70%) and testing (30%) sets. Model performance was assessed using accuracy, sensitivity, specificity, and area under the ROC curve. Five-fold cross-validation was used to validate the models.

The best-performing model was selected based on overall accuracy and deployed as a portable application using Streamlit (Streamlit Inc., San Francisco, CA, USA). Internal validation was performed using a hold-out subset of the original data, while external validation utilized an independent dataset from another hospital. All statistical tests were two-sided, and a p-value < 0.05 was considered statistically significant.

Results

Characteristics of Respondents

In this study, 369 colorectal cancer patients visited the outpatient polyclinic and inpatient installation of Digestive Surgery at Dr. Wahidin Sudirohusodo Central General Hospital, Makassar, South Sulawesi. The characteristics of the study subjects can be seen in Table 1.

The mean age of the patients was 54.7 ± 14.1 years. The majority of the study subjects were male (57.%), had cancer location in the rectum (54.2%), diagnosed at stage 3 (45%), had tumor size T3 (51.7%), had KGB enlargement (67.5%), and had no metastasis (76.6%). Since diagnosis until now, there have been 16 cases (4.4%) of loss to follow-up. A total of 58.5% of patients were still alive from the initial diagnosis until the study was conducted, with a mean survival time of 27.3 ± 1.4 months.

Table 1. Study Participant Characteristics

Variable	Number (n)	Percentage (%)
Gender		
Male	213	57.7
Female	156	42.3
Location		
Colon Ascendens	60	16.3
Colon transversum	16	4.3
Colon descendens	23	6.2
Colon sigmoid	70	19.0
Rektum	200	54.2
Stadium		
1	23	6.3
2	95	25.7
3	166	45.0
4	85	23.0
Tumor (T)		
T2	146	39.6
T3	191	51.7
T4	32	8.7
Lymph Nodes (N)		
N0	120	32.5
N1	249	67.5
Metastasis (M)		
M0	283	76.7
M1	86	23.3
Status		
Live	216	58.5
Death	137	37.1
Loss to follow up	16	4.4
Management		
Definitive Operation	263	71.3
Diversion Operation	98	26.5
No Operation	8	2.2
Adjuvant Chemotherapy		
Yes	184	70
No	79	30
Neoadjuvant Chemotherapy		
Yes	24	4.1
No	74	75.9
Recurrency		
Yes	16	6.5
No	168	93.5

A total of 263 out of 369 patients (71.3%) with colorectal cancer underwent definitive surgery, 8 patients (2.2%) had preoperative progressivity, and 98 patients (26.5%) underwent only diversionary surgery (stoma or bypass). Of the 263 patients who underwent definitive surgery, 79 patients (30%) did not continue adjuvant chemotherapy, while the remaining 184 patients (70%) continued adjuvant chemotherapy with Folfiri (12.5%); Folfox (26.6%); and Capeox (60.9%) regimens. A total of 16 out of 184 patients (6.5%) who underwent adjuvant chemotherapy experienced recurrence.

Of the 98 patients who underwent diversion surgery (stoma or bypass), the majority 70 patients (71.4%) did not undergo chemotherapy, 4 patients (4.1%) were lost to follow-up, and only 24 patients (24.5%) underwent chemotherapy with Capeox (19 patients), Folfox (4 patients), and Folfiri (1 patient) regimens. Response to neoadjuvant chemotherapy was shown to be good with 17 patients (70.8%) having a partial response and being able to proceed to definitive surgery while 7 patients (29.2%) had a progressive response or stable disease.

Platelet Indices on Colorectal Cancer Staging

This study highlights the potential of platelet indices as valuable biomarkers in understanding colorectal cancer progression. This study employed non-parametric statistical analysis due to the non-normal distribution of the data, specifically utilizing the Kruskal-Wallis test to assess these relationships (Table 2). The investigation revealed a striking trend: as the stage of colorectal cancer advanced, the platelet count (PC) increased significantly. Similarly, MPV demonstrated a positive correlation with cancer progression. PDW also increased with advancing stages, indicating variability in platelet size. The analysis showed that PCT rose alongside the cancer stage. In contrast to the other indices, the MPV/PC ratio exhibited a decrease as the disease progressed.

Machine Learning-Based Predictor Scoring System for Colorectal Cancer Staging

In developing a prediction model based on platelet indices for colorectal cancer staging, a bivariate test was conducted to examine the relationships between gender, tumor location, and age with colorectal cancer stage. The statistical results indicated no significant correlation between these three variables and cancer stage, allowing the continuation of the prediction model using platelet indices.

In this context, since the statistical power of Platelet Count (PC) and Mean Platelet Volume (MPV) was found to be higher individually than that of the MPV/PC ratio, the authors chose to include only PC and MPV as separate variables, excluding the ratio. The results from the Kruskal-Wallis test indicated that the platelet indices contributing most significantly to predicting colorectal cancer stage were ranked as follows: PC, Platelet Distribution Width (PDW), Plateletcrit (PCT), and lastly MPV.

In this study, data that did not follow a normal distribution resulted in a low prediction model accuracy of 51.1% when using simple regression. To enhance predictive performance, a machine learning-based model was designed using MATLAB. The results indicated that the model utilizing the Support Vector Machine (SVM) algorithm achieved the best performance, with an accuracy of 82.9%. This was closely followed by the K-Nearest Neighbors (KNN) algorithm at 82.7%, and the Neural Network model at 81.5%. The Naive Bayes

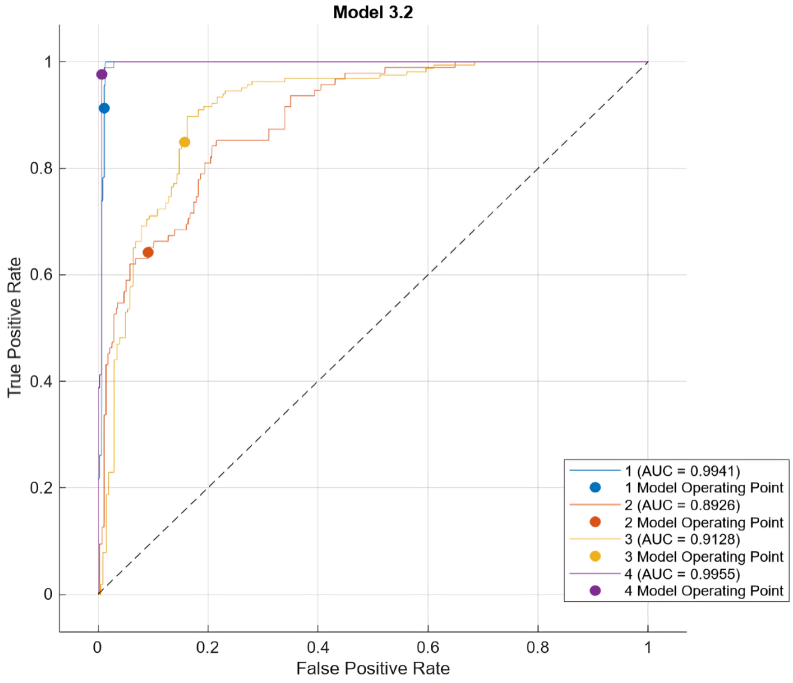

Figure 1. ROC Curve of the Colorectal Cancer Staging Prediction Model Based on SVM

classifier reached an accuracy of 80.5%, while the logistic regression model performed the worst, with an accuracy of only 51.5%.

The SVM-based prediction model demonstrated strong metrics in its ROC curve analysis, detailed as follows (Figure 1):

Stage 1: AUC 100%; True Positive Rate (TPR) 93%; Positive Predictive Value (PPV) 84%

Stage 2: AUC 96.5%; TPR 64.2%; PPV 70.9%

Stage 3: AUC 93%; TPR 84.9%; PPV 81.5%

Stage 4: AUC 98.6%; TPR 97.6%; PPV 97.6%

These findings suggest that machine learning algorithms, particularly SVM, can significantly enhance

Table 2. The Relationship between Platelet Index and the Stage of Colorectal Cancer Patients at Dr. Wahidin Sudirohusodo Central General Hospital

Variable	n	PC	PDW	MPV	PCT	Ratio MPV/PC
		Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)
Stadium						
1	23	289 (102)	8.2 (1.8)	8.1 (1.5)	0 (0)	0.034 (0.002)
2	95	310 (39)	8.2 (0.2)	8.2 (0.2)	0 (0)	0.027 (0.003)
3	166	440 (84)	8.8 (0.4)	8.7 (0.4)	0.2 (0.1)	0.019 (0.002)
4	85	665 (149)	10.6 (1.1)	10.5 (1.4)	0.4 (0.1)	0.016 (0.001)
p	369	<0.05	<0.05	<0.05	<0.05	<0.05
Tumor						
T2	32	225 (25,5)	7.8 (0.6)	7.2 (0.4)	0 (0)	0.032 (0.001)
T3	191	347 (39,5)	8.5 (0.4)	8.5 (0.4)	0.1 (0.1)	0.024 (0.004)
T4	146	582,5 (65,3)	9.9 (0.9)	9.9 (1)	0.3 (0.1)	0.0.17 (0.001)
p	369	<0.05	<0.05	<0.05	<0.05	<0.05
Lymph Node						
No	120	290 (40,5)	8.1 (0.4)	8.1 (0.6)	0 (0)	0.028 (0.004)
Yes	249	521 (132)	9.2 (0.5)	9.2 (0.7)	0.2 (0.1)	0.018 (0.002)
p	369	<0.05	<0.05	<0.05	<0.05	<0.05
Metastasis						
No	283	359 (52)	8.5 (0.4)	8.5 (0.4)	0.1 (0)	0.024 (0.005)
Yes	86	662 (51,5)	10.6 (0.4)	10.5 (0.3)	0.4 (0)	0.016 (0.001)
p	369	<0.05	<0.05	<0.05	<0.05	<0.05

PC, Platelet Concentrate; PDW, Platelet Distribution Width; MPV, Mean Platelet Volume; PCT, packed cell volume; MPV/PC, Mean Platelet Volume/ Platelet Concentrate; IQR, Inter-Quartile Range.

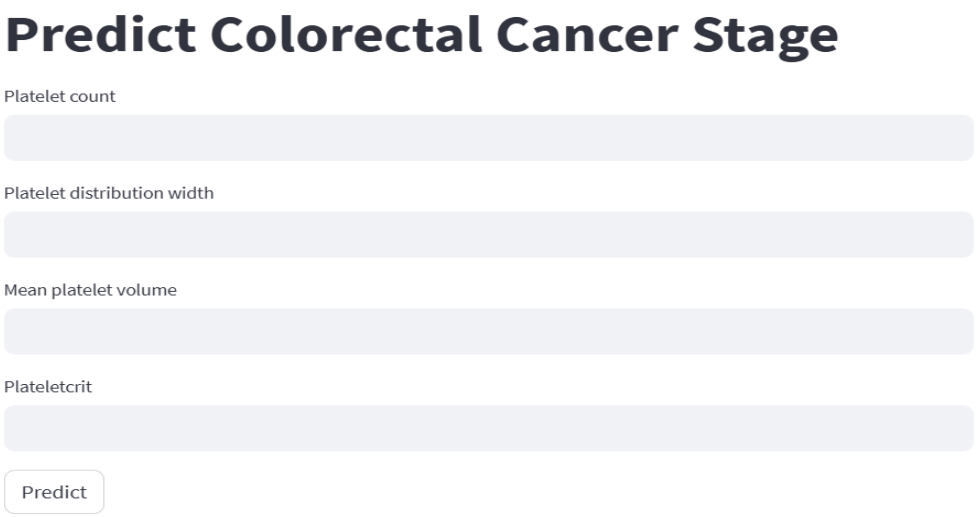

Figure 2. Application of the SVM-Based Predictive Model for Colorectal Cancer Staging from Platelet Indices

Table 3. Differences in Platelet Indices in Colorectal Cancer Patients Before and After Definitive Surgery

Platelet Indexes	N	Median (IQR) before	Median (IQR) After	p
Definitive Operation				
PC	263	363 (56)	301 (50)	<0.05
PDW	263	8.5 (0.4)	8.1 (0.5)	<0.05
MPV	263	8.5 (0.4)	8.1 (0.4)	<0.05
PCT	263	0.1 (0.1)	0 (0)	<0.05
Adjuvant Chemotherapy				
PC	184	288.5 (55.2)	214.5 (35.5)	<0.05
PDW	184	8.1 (0.6)	7.2 (0.5)	<0.05
MPV	184	8.1 (0.4)	7.8 (0.7)	<0.05
PCT	184	0 (0)	0 (0)	<0.05
Recurrency				
PC	184	278.5 (62.8)	551.5 (89)	<0.05
PDW	184	8.2 (0.7)	10.5 (0.3)	<0.05
MPV	184	8.15 (0.45)	10.2 (0.4)	<0.05
PCT	184	0 (0.05)	0 (0)	<0.05

the accuracy of colorectal cancer staging predictions compared to traditional regression methods.

The predictive model was further developed using the Python programming language to establish both internal and external validation, employing a train-test split of 70:30. The internal validation yielded an accuracy of 79.2%, while the external validation reached 89.2% (Table 3).

This model has been deployed as a portable application on Streamlit, accessible at the following link: https://trombositbaru-uekiugszspno4u4wxqqpio.streamlit.app/ (see Figure 2). This deployment allows users to interact with the predictive model easily, enhancing its accessibility for real-world applications.

Furthermore, the researchers were interested in evaluating the differences in platelet indices during the progression of colorectal cancer, specifically before and after definitive surgery, after adjuvant chemotherapy, and in cases of recurrence. The results from the Wilcoxon test indicated that all platelet indices significantly decreased after definitive surgery compared to measurements taken before the operation ($p<0.05$), decreased after adjuvant chemotherapy compared to pre-chemotherapy levels ($p<0.05$), and significantly increased in cases of recurrence ($p<0.05$).

Discussion

The platelet index comprises parameters that reflect both the quantity and size of platelets, including platelet count (PC), mean platelet volume (MPV), platelet distribution width (PDW), and plateletcrit (PCT). Notably, the MPV/PC ratio has proven valuable in cancer diagnosis and prognosis [13, 14]. However, in Makassar City, research exploring the relationship between the platelet index and the clinicopathology of colorectal cancer remains absent, highlighting a critical gap that this study aims to address.

Relationship Between PC and Colorectal Cancer Stage

Platelet count (PC) refers to the number of platelets in the blood and can be assessed using routine automatic hematology machines [15]. Importantly, the interaction between platelets and cancer is reciprocal; while platelets stimulate cancer development, carcinogenesis also alters platelet characteristics and functions. Studies have shown that malignant tumors can increase both the number and activity of platelets, thereby promoting cancer progression [16]. Furthermore, platelets play a crucial role in supporting tumor growth and metastasis by releasing various growth factors that facilitate angiogenesis and metastatic spread [17].

This study demonstrates a significant relationship between platelet count (PC) and the stage of colorectal cancer, showing that platelet counts rise with disease progression, tumor status, lymph node involvement, and metastasis. This aligns with the findings of Ramjeesingh et al. [18], which indicated that thrombocytosis significantly increased at stage 4. Furthermore, Lin et al. [19] found that elevated platelet counts correlated significantly with lymph node status (p=0.016), distant metastasis (p=0.014), and cancer stage (p=0.014).

In this study, thrombocytosis (cut-off 400 x 10^9/L) was predominantly found in stages 3 and 4, aligning with Guo et al. [20]'s research reported thrombocytosis in 75% of advanced-stage patients (stage III and IV). No patients with stage I cancer exhibited increased platelet counts at diagnosis [20]. Thrombocytosis differentiates T4 status, lymph node involvement (N1), and the presence of metastasis (M1). The group with metastasis (M1) showed very high platelet levels (reaching 662). Excessive platelet counts significantly increase the risk of metastasis at every cancer stage and serve as a poor prognostic indicator, particularly in stomach, lung, and kidney cancers [21].

Relationship Between MPV and Colorectal Cancer Stage

Mean Platelet Volume (MPV) is defined as the ratio of plateletcrit (PCT) to platelet count (PC) and reflects the average volume of circulating platelets [22]. Importantly, in pathological conditions like inflammation, increased megakaryopoiesis and thrombopoiesis elevate both PC and MPV, along with other platelet indices [23].

This study shows a significant relationship between MPV and the stage of colorectal cancer patients, where MPV values increase with stage progression. However, the median MPV at each stage did not exceed the normal laboratory limits (6.5-11). This aligns with the retrospective study by Li et al. [24], which demonstrated a positive correlation between MPV values and TNM stage in colorectal cancer [25].

The carcinogenesis process is linked to a hyper-inflammatory state with the secretion of pro-coagulant and pro-inflammatory factors. Cytokines, especially IL-6, trigger thrombopoietin formation, leading to increased ploidy of megakaryocyte nuclei and enhanced cytoplasmic volume. Consequently, both platelet count and MPV increase [26].

Larger platelets are more involved in tissue infiltration. They aggregate more rapidly with collagen and have increased granule secretion, more glycoprotein Ib and IIb/IIIa receptors, and higher levels of thromboxane A2. β-Thromboglobulin and platelet factor 4, specific proteins released from platelet α-granules, all play roles in tumor growth progression [27].

Relationship Between PDW and Colorectal Cancer Stage

Platelet Distribution Width (PDW) measures the volume dispersion of platelets and is defined as the coefficient of variation of platelet volume, reflecting changes in platelet size and heterogeneity [28, 29]. Interestingly, changes in PDW across various cancer types are inconsistent; for example, PDW increases in gastric and lung cancers but decreases in thyroid and breast cancers [30].

In this study, we found a significant relationship between PDW and the stage of colorectal cancer, with PDW values rising as the stage progresses. However, it is noteworthy that the median PDW at each stage remained within normal laboratory limits [10-18]. The mechanism behind increased PDW in malignancies likely involves interactions between the tumor microenvironment and activated platelets. Tumor cells secrete cytokines such as interleukin-6 (IL-6), granulocyte colony-stimulating factor (G-CSF), and macrophage colony-stimulating factor (M-CSF), which stimulate megakaryopoiesis. Activated platelets contribute to a hypercoagulable state, allowing them to envelop tumor cells and evade the host immune response, resulting in elevated PDW counts. Conversely, the specific mechanisms explaining reduced PDW in certain cancers remain unclear [31].

Relationship Between PCT and Colorectal Cancer Stage

Plateletcrit (PCT) represents the volume occupied by platelets in the blood as a percentage [32, 33]. This study reveals a significant relationship between PCT and the stage of colorectal cancer, with PCT values rising as the stage progresses. Notably, the median PCT at each stage remained within normal laboratory limits (0.15-0.5). Supporting this, Zhu and Cao [34] reported a significant increase in PCT with advancing colorectal cancer stages ($p<0.001$), further emphasizing its potential role as a biomarker in cancer progression.

Plateletcrit (PCT) is the volume occupied by platelets in the blood as a percentage [32, 33]. PCT values vary between patients and healthy subjects across different cancer types. This study indicates a significant relationship between PCT and the stage of colorectal cancer patients, where PCT values increase with stage progression. However, the median PCT at each stage did not exceed the normal laboratory limits (0.15-0.5). Zhu et al. [34] reported that PCT significantly increased with the advancement of colorectal cancer stages ($p<0.001$).

Relationship Between MPV/PC Ratio and Colorectal Cancer Stage

The MPV/PC ratio is calculated as MPV measured in 10−15 L divided by the absolute platelet count measured in $\times 10^9$/L. Some researchers argue that the MPV to PC ratio should be interpreted as a ratio rather than used in isolation.

This study shows a significant relationship between

the MPV/PC ratio and the stage of colorectal cancer patients, where the MPV/PC ratio decreases with increasing stage. Wu et al. [2019] found that the MPV/PC value significantly differentiates colorectal cancer stages (p=0.021). Zhang et al. [35] also demonstrated a statistically significant relationship between the MPV/PC ratio and T and N stages (p=0.0007 and p=0.0079).

Prediction Model for Colorectal Cancer Staging Based on Platelet Index

The prediction models for colorectal cancer staging are still limited, as most models have been developed to predict survival rather than staging. Only three studies have focused on developing prediction models specifically for colorectal cancer stages.

Gupta et al. [36] created a model to predict T-stage colorectal cancer using medical records from 4,021 patients in Taiwan. This model considered variables such as body mass index, family history, age, gender, history of hypertension, diabetes, smoking habits, alcohol consumption, CEA levels, albumin, creatinine, and leukocytes. The machine-learning method they used achieved an accuracy of 90% with a random forest approach [36]. In particular, they developed a model to predict only the T stage of colorectal cancer, utilizing variables such as age, CEA, CA19-9, tumor location, and characteristics from CT scans, achieving an accuracy of 87%.

Lu et al. [37] developed a deep learning system for predicting colorectal cancer stages and RAS mutations. This study used abdominal CT scans with contrast, combined with CEA levels, age, and gender. The neural network model achieved a validation score of 98% for staging and 95% for RAS mutations. However, a limitation of this model is its reliance on tumor biomarkers that may not be available in all healthcare facilities [37].

Routine automatic hematology testing is one of the most frequently performed tests in clinical laboratories. Its relatively low cost, high reproducibility, and flexibility make it suitable for further development as a prediction model for colorectal cancer staging. To date, there has been no prediction model for colorectal cancer staging based on platelet index data. The machine learning model designed in this study provides good predictive value compared to regression prediction models, as the data in this study was not normally distributed. The SVM model built from PC, PDW, MPV, and PCT achieved an internal validation score of 79.2% and an external validation score of 89.2%.

In conclusion, this study highlights the platelet index as an exciting predictor for staging colorectal cancer, offering a groundbreaking opportunity for clinical practice. The developed predictive model not only enhances diagnostic accuracy but also fosters a deeper connection to patient care.

To fully realize this potential, it's crucial to discuss how clinicians can integrate this tool into their workflows, addressing barriers such as training and technology access. Beyond colorectal cancer, these findings could reshape public health policies and inform screening guidelines, promoting earlier detection and intervention.

Ultimately, this non-invasive approach represents a significant advancement in cancer management, appealing to both clinicians and patients. By embracing this innovation, we can pave the way for a more personalized and effective strategy in tackling colorectal cancer and improving public health outcomes overall.

Author Contribution Statement

CA is responsible for Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Roles/Writing - original draft; Writing - review & editing. REL, SRL, IL, W, MAF, M, SS, MIK, JAU, and ES are responsible for conceptualization, methodology, and supervision.

Acknowledgements

The study protocol was approved by the Ethics Committee of Wahidin Sudirohusodo Makasar General Hospital and the Institutional Review Board of the Hasanuddin University Faculty of Medicine (Makassar, Indonesia; the clinical trial registration number: UH24010063).

Availability of data

The data is available as request to authors.

Ethical Clearance

The study protocol was approved by the Ethics Committee of Wahidin Sudirohusodo Makasar General Hospital and the Institutional Review Board of the Hasanuddin University Faculty of Medicine (Makassar, Indonesia; the clinical trial registration number: UH24010063).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors. Przeglad gastroenterologiczny. 2019;14(2):89-103. https://doi.org/10.5114/pg.2018.81072.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424. https://doi.org/10.3322/caac.21492.
3. WHO. Indonesia, source: Globocan 2020. IARC. 2020.
4. Labeda I, Lusikooy RE, Mappincara, Dani MI, Sampetoding S, Kusuma MI, et al. Colorectal cancer survival rates in makassar, eastern indonesia: A retrospective cohort study. Ann Med Surg (Lond). 2022;74:103211. https://doi.org/10.1016/j.amsu.2021.103211.
5. Arsyad A, Lusikooy R, Syarifuddin E, Rahardjo W, Labeda I, Mappincara, et al. Analysis of the prognostic factors

affecting 5-year colorectal cancer survival rates in makassar, eastern indonesia: A retrospective cohort study. Gazz Med Ital. 2022;181:1-8. https://doi.org/10.23736/S0393-3660.22.04923-3.
6. Yang Y, Wang Y, Wang Z. Construction of a new clinical staging system for colorectal cancer based on the lymph node ratio: A validation study. Front Surg. 2022;9:929576. https://doi.org/10.3389/fsurg.2022.929576.
7. Smith RA, Oeffinger KC. The importance of cancer screening. Med Clin North Am. 2020;104(6):919-38. https://doi.org/10.1016/j.mcna.2020.08.008.
8. Buergy D, Wenz F, Groden C, Brockmann MA. Tumor-platelet interaction in solid tumors. Int J Cancer. 2012;130(12):2747-60. https://doi.org/10.1002/ijc.27441.
9. Haemmerle M, Stone RL, Menter DG, Afshar-Kharghan V, Sood AK. The platelet lifeline to cancer: Challenges and opportunities. Cancer Cell. 2018;33(6):965-83. https://doi.org/10.1016/j.ccell.2018.03.002.
10. Golebiewska EM, Poole AW. Platelet secretion: From haemostasis to wound healing and beyond. Blood Rev. 2015;29(3):153-62. https://doi.org/10.1016/j.blre.2014.10.003.
11. Zhang Z, Xu X, Ni H, Deng H. Platelet indices are novel predictors of hospital mortality in intensive care unit patients. J Crit Care. 2014;29(5):885.e1-6. https://doi.org/10.1016/j.jcrc.2014.04.020.
12. Karagöz B, Sücüllü İ, Sayan Ö, Bilgi O, Tuncel T, Filiz Aİ, et al. Platelet indices in patients with colorectal cancer. Cent Eur J Med. 2010;5(3):365-8. https://doi.org/10.2478/s11536-009-0077-7.
13. Zhang X, Cui MM, Fu S, Li LL, Liu YS, Liu ZP, et al. Platelet distribution width correlates with prognosis of gastric cancer. Oncotarget. 2017;8(12):20213-9. https://doi.org/10.18632/oncotarget.15561.
14. Oncel M, Kiyici A, Oncel M, Sunam GS, Sahin E, Adam B. Evaluation of platelet indices in lung cancer patients. Asian Pac J Cancer Prev. 2015;16(17):7599-602. https://doi.org/10.7314/apjcp.2015.16.17.7599.
15. Giannakeas V, Kotsopoulos J, Cheung MC, Rosella L, Brooks JD, Lipscombe L, et al. Analysis of platelet count and new cancer diagnosis over a 10-year period. JAMA Network Open. 2022;5(1):e2141633-e. https://doi.org/10.1001/jamanetworkopen.2021.41633.
16. Lin RJ, Afshar-Kharghan V, Schafer AI. Paraneoplastic thrombocytosis: The secrets of tumor self-promotion. Blood. 2014;124(2):184-7. https://doi.org/10.1182/blood-2014-03-562538.
17. Goubran HA, Burnouf T, Radosevic M, El-Ekiaby M. The platelet-cancer loop. Eur J Intern Med. 2013;24(5):393-400. https://doi.org/10.1016/j.ejim.2013.01.017.
18. Ramjeesingh R, Jones A, Orr C, Bricks CS, Hopman WM, Hammad N. Thrombocytosis as a predictor of poor prognosis in colorectal cancer patients. J Clin Oncol. 2016;34(4_suppl):540-. https://doi.org/10.1200/jco.2016.34.4_suppl.540.
19. Lin MS, Huang JX, Zhu J, Shen HZ. Elevation of platelet count in patients with colorectal cancer predicts tendency to metastases and poor prognosis. Hepatogastroenterology. 2012;59(118):1687-90. https://doi.org/10.5754/hge12277.
20. Guo T, Krzystanek M, Szallasi Z, Szallasi A. Thrombocytosis portends adverse prognostic significance in patients with stage ii colorectal carcinoma. F1000Res. 2014;3:180. https://doi.org/10.12688/f1000research.4856.2.
21. Caine GJ, Stonelake PS, Lip GY, Kehoe ST. The hypercoagulable state of malignancy: Pathogenesis and current debate. Neoplasia (New York, NY). 2002;4(6):465-73. https://doi.org/10.1038/sj.neo.7900263.
22. Ntolios P, Papanas N, Nena E, Boglou P, Koulelidis A, Tzouvelekis A, et al. Mean platelet volume as a surrogate marker for platelet activation in patients with idiopathic pulmonary fibrosis. Clin Appl Thromb Hemost. 2016;22(4):346-50. https://doi.org/10.1177/1076029615618023.
23. Chu SG, Becker RC, Berger PB, Bhatt DL, Eikelboom JW, Konkle B, et al. Mean platelet volume as a predictor of cardiovascular risk: A systematic review and meta-analysis. J Thromb Haemost. 2010;8(1):148-56. https://doi.org/10.1111/j.1538-7836.2009.03584.x.
24. Li R, Ren M, Chen N, Luo M, Deng X, Xia J, et al. Presence of intratumoral platelets is associated with tumor vessel structure and metastasis. BMC Cancer. 2014;14:167. https://doi.org/10.1186/1471-2407-14-167.
25. Li JY, Li Y, Jiang Z, Wang RT, Wang XS. Elevated mean platelet volume is associated with presence of colon cancer. Asian Pac J Cancer Prev. 2014;15(23):10501-4. https://doi.org/10.7314/apjcp.2014.15.23.10501.
26. Senchenkova EY, Komoto S, Russell J, Almeida-Paula LD, Yan LS, Zhang S, et al. Interleukin-6 mediates the platelet abnormalities and thrombogenesis associated with experimental colitis. The Am J Pathol. 2013;183(1):173-81. https://doi.org/10.1016/j.ajpath.2013.03.014.
27. Sobolewska A, Włodarczyk M, Stec-Michalska K, Fichna J, Wiśniewska-Jarosińska M. Mean platelet volume in crohn's disease patients predicts sustained response to a 52-week infliximab therapy: A pilot study. Dig Dis Sci. 2016;61(2):542-9. https://doi.org/10.1007/s10620-015-3894-3.
28. Sachdev R, Tiwari AK, Goel S, Raina V, Sethi M. Establishing biological reference intervals for novel platelet parameters (immature platelet fraction, high immature platelet fraction, platelet distribution width, platelet large cell ratio, platelet-x, plateletcrit, and platelet distribution width) and their correlations among each other. Indian J Pathol Microbiol. 2014;57(2):231-5. https://doi.org/10.4103/0377-4929.134676.
29. Maluf CB, Barreto SM, Vidigal PG. Standardization and reference intervals of platelet volume indices: Insight from the brazilian longitudinal study of adult health (elsa-brasil). Platelets. 2015;26(5):413-20. https://doi.org/10.3109/09537104.2014.942620.
30. Cui MM, Li N, Liu X, Yun ZY, Niu Y, Zhang Y, et al. Platelet distribution width correlates with prognosis of non-small cell lung cancer. Sci Rep. 2017;7(1):3456. https://doi.org/10.1038/s41598-017-03772-z.
31. Siddeek RAT, Gupta A, Gupta S, Goyal B, Gupta AK, Agrawal S, et al. Evaluation of platelet distribution width as novel biomarker in gall bladder cancer. J Carcinog. 2020;19:5. https://doi.org/10.4103/jcar.JCar_12_20.
32. Wiwanitkit V. Plateletcrit, mean platelet volume, platelet distribution width: Its expected values and correlation with parallel red blood cell parameters. Clin Appl Thromb Hemost. 2004;10(2):175-8. https://doi.org/10.1177/107602960401000208.
33. Budak YU, Polat M, Huysal K. The use of platelet indices, plateletcrit, mean platelet volume and platelet distribution width in emergency non-traumatic abdominal surgery: A systematic review. Biochem Med (Zagreb). 2016;26(2):178-93. https://doi.org/10.11613/bm.2016.020.
34. Zhu X, Cao Y, Lu P, Kang Y, Lin Z, Hao T, et al. Evaluation of platelet indices as diagnostic biomarkers for colorectal cancer. Sci Rep. 2018;8(1):11814. https://doi.org/10.1038/s41598-018-29293-x.
35. Zhang H, Lin F, Wang Z. Mean platelet volume/platelet count ratio in combination with tumor markers in colorectal cancer: A retrospective clinical study. BMC Cancer. 2023;23(1):124.

https://doi.org/10.1186/s12885-023-10585-z.
36. Gupta P, Chiang SF, Sahoo PK, Mohapatra SK, You JF, Onthoni DD, et al. Prediction of colon cancer stages and survival period with machine learning approach. Cancers (Basel). 2019;11(12). https://doi.org/10.3390/cancers11122007.
37. Lu N, Guan X, Zhu J, Li Y, Zhang J. A contrast-enhanced ct-based deep learning system for preoperative prediction of colorectal cancer staging and ras mutation. Cancers. 2023;15(18):4497. https://doi.org/10.3390/cancers15184497