

Received: 2020.11.16

Accepted: 2021.05.10

Available online: 2021.06.02

Published: 2021.08.27

# A Prognostic Model of Gastrointestinal Diffuse Large B Cell Lymphoma

Authors' Contribution:

Study Design A  
Data Collection B  
Statistical Analysis C  
Data Interpretation D  
Manuscript Preparation E  
Literature Search F  
Funds Collection G

ACDEF 1 **Fan Gao**  
BCD 2 **Zhan-fang Wang**  
DF 1 **Lei Tian**  
BC 1 **Fei Dong**  
BD 1 **Jing Wang**  
BD 1 **Hong-mei Jing**  
AG 1 **Xiao-yan Ke**

1 Department of Hematology, Peking University Third Hospital, Beijing, PR China  
2 Department of Hematology, Pingdingshan First People's Hospital, Pingdingshan, Henan, PR China

**Corresponding Author:** Xiao-yan Ke, e-mail: xykbysy@163.com  
**Financial support:** Departmental sources

**Background:** The digestive tract is the most common site of extranodal involvement in diffuse large B cell lymphoma (DLBCL) and its prognostic evaluation is different from that of ordinary DLBCL. Currently, for gastrointestinal lymphoma, in addition to the Ann Arbor staging system, the Lugano and the TNM staging systems are commonly used. However, there is no effective prognostic model to identify poor prognosis in patients with localized gastrointestinal diffuse large B cell lymphoma (GI-DLBCL).


**Material/Methods:** This study included 82 patients with GI-DLBCL that had a median follow-up of 75 months, and developed a model (HLAMA) with 5 variables: hemoglobin, age, lactate dehydrogenase (LDH), serum albumin, and the maximum intra-abdominal lesion diameter (MIALD). The specific indicators are: HGB <105 g/L (2 points); LDH ≥300 U/L; age ≥75 years, ALB <38 g/L, MIALD ≥4 cm (each scoring 1 point). We also developed a simplified model, which includes only 3 variables (HGB, LDH, and age).

**Results:** HLAMA model and the simplified model both demonstrated good ability to predict prognosis of patients with GI-DLBCL ( $P < 0.001$ ), performing better than the IPI score as it could distinguish low-risk groups in relatively elderly patients (60-75 years old).

**Conclusions:** This study established a prognostic model for diffuse large B cell lymphoma of the gastrointestinal tract. Both the HLAMA model and its simplified version are similar to the IPI score, but could be considered better as they can provide a simpler and more accurate prognostic assessment in patients with GI-DLBCL. For patients with localized GI-DLBCL, our model could distinguish high-risk patients.

**Keywords:** Gastric Mucosa • Lymphoma, Large B-Cell, Diffuse • Primary Graft Dysfunction • Prognosis • Tumor Stem Cell Assay

Full-text PDF: <https://www.medscimonit.com/abstract/index/idArt/929898>

 3725

 4

 6

 31



## Background

The gastrointestinal tract is the most common site of extranodal involvement in diffuse large B cell lymphoma (DLBCL), accounting for approximately one-third of cases of primary extranodal DLBCL [1]. The Ann Arbor staging system [2,3] is not easily applied to gastrointestinal tract lymphomas and, although alternative staging systems have been proposed, staging gastrointestinal non-Hodgkin lymphoma (PG-NHL) is still controversial [4]. Alternative staging systems include the Lugano and the TNM staging systems. In the Lugano staging system, prognosis is based on gastrointestinal involvement, especially distant lymph node involvement. The high-risk group, Lugano stage IV, specifically refers to patients with diffuse extranodal involvement or concomitant supracondylar lymph node involvement [5]. This system has a good prognostic utility, but like the Ann Arbor staging system, it is difficult to effectively predict the prognosis in patients with DLBCL that have limited gastrointestinal tract involvement.

The TNM staging system derives from the staging system of solid tumors and is evaluated by primary tumor characteristics (T), the presence or absence of regional lymph node involvement (N), and the presence or absence of distant metastases (M). Based on a retrospective analysis of 101 patients with primary gastrointestinal lymphoma, Chang et al [6] concluded that the TNM staging system was better than the Ann Arbor and the Lugano staging systems in the prognostic evaluation of primary gastrointestinal lymphoma. However, because patients with gastrointestinal diffuse large B cell lymphoma do not always undergo surgery and deep biopsy, the assessment of the depth of invasion is clinically limited. Although endoscopic ultrasound (EUS) can determine the depth of the invasion, additional tests increase the cost of treatment and are not carried out by many health centers.

The International Prognostic Index (IPI) is one of the most important prognostic systems in non-Hodgkin lymphoma, with accurate prognosis in several subtypes. In 2015, researchers found that the IPI score not only clearly separated patients with primary gastric diffuse large B cell lymphoma (PG-DLBCL) into different risk groups, but also separated early-stage patients, classified according to the Lugano system, into groups with distinct prognosis [7]. However, survival of patients with the same IPI score varied significantly, indicating that there is significant residual heterogeneity in each IPI category [8]. Moreover, the original IPI score is often only available in 3 dimensions (age, lactate dehydrogenase, and performance status) in patients with limited-stage GI-DLBCL [9]. This will inevitably have an impact on the effectiveness of the model for this group of patients.

We aimed to evaluate the prognostic significance of the IPI score in gastrointestinal DLBCL based on clinical manifestations and laboratory tests. Specifically, we aimed to identify patients with limited-stage PG-DLBCL that have a poor prognosis. We expect that a new prognostic model can be established to more easily and accurately evaluate the prognosis of DLBCL of the gastrointestinal tract.

## Material and Methods

### Data Set Summary

This study included 82 DLBCL patients with gastrointestinal involvement who were treated at the Peking University Third Hospital from March 2007 to March 2017. All patients were diagnosed by gastrointestinal endoscopy or diagnostic biopsy, and biopsy samples from recurrence cases were ruled out. The Hans algorithm was used to perform immunohistochemical classification of diffuse large B cell lymphoma. The median follow-up time was 75 months. Among the 82 patients, the male-female ratio was 1: 2. The median age was 63.5 years and 15 patients (18.3%) were over age 75. The diagnosis of primary gastrointestinal lymphoma is based on Dawson et al [10]. There were 38 cases (46.3%) of primary gastrointestinal lymphoma, 42 cases (51.2%) with stomach involvement, 12 cases (14.6%) with small intestine involvement (including the duodenum), 14 cases (17.1%) with ileocecal involvement, and 4 cases (4.9%) with colon involvement. Rectal involvement was rare, with only 1 patient and a total of 9 patients (11%) showing multiple site involvement (2 or more sites).

We grouped patients based on their initial symptoms and endoscopic findings. Regarding symptoms, patients could be: (1) asymptomatic; (2) show specific symptoms of the digestive system, which included hematemesis, hematochezia, melena, and obstipation (ie, inability to pass flatus or stool); and (3) show non-specific digestive symptoms, which included abdominal pain, bloating, anorexia, nausea, and vomiting. Regarding endoscopic findings, patients were classified according to: (1) types of ulcers: no ulcer, isolated simple ulcers, giant ulcers, and multiple ulcers; and (2) types of lesions: edema or benign ulcers (no obvious bulging lesions or lumps, only edema or ulcer lesions), thickening and diffuse bulges (with microscopically visible bulges or thickenings), lumps, or polyp-type lesions.

We used abdominal computed tomography (CT) or positron emission tomography (PET)-CT to develop a new index in which the maximum intra-abdominal measurable lesion diameter (MIALD) in the abdominal cavity, including gastrointestinal lesions and affected lymph nodes, was measured. We also analyzed the patient's pathological subtypes and immunohistochemical markers (Bcl-2 and Bcl-6). GCB and ABC

were distinguished according to the Hans algorithm and the Tally method [11,12]. Laboratory indicators used in this study included: HGB (hemoglobin, 130-175 g/L), LDH (lactic dehydrogenase, 120-250 U/L), ALB (albumin, 40-55g/L), and  $\beta 2$  microglobulin ( $\beta 2$ -MG, 1.0-3.0 mg/L). These indicators were measured at the patient's initial diagnosis. HGB refers to the value when no blood transfusion treatment was received at the initial diagnosis. The number and proportion of patients in each group can be seen in **Table 1**.

### Variable Selection and Statistical Methods

For all quantitative data, based on the 3-year overall survival (OS) rate, we used the pROC package [13] (version 1.15.3) in R software to draw the ROC curve and select the appropriate cut-off values. For all categorical variables, comparisons between groups were made based on the 3-year survival rate and *P* values were obtained by chi-square tests and Fisher's exact tests. Survival analysis was performed on each group and included age, sex, Ann Arbor stage, IPI score, onset site, and complications. All *P* values were obtained by the log-rank test. We used a random forest model [14] (randomForest package in R software, version 4.6-14) and a COX regression model to screen important variables and construct our prognostic models. Statistical analyses were carried out with SPSS 25.0 and R software (version 3.5.3).

## Results

### Three-year Survival Rate

We calculated the 3-year survival rate of all groups and performed a chi-square test to assess the differences between variables (**Table 1**). The 3-year survival rates for males and females were similar, and there was no significant difference in survival between groups with different sites involved. The clinical symptoms, pre-treatment complications, pathological subtypes, and BCL-2 and BCL-6 expression did not differ between groups. Additionally, the endoscopic findings did not affect the 3-year survival rate of patients. The significant variables that affected patient survival rate were: age ( $P=0.01$ , HR=2.1), primary gastrointestinal lymphoma ( $P<0.05$ , HR=2.84), post-treatment complications ( $P<0.05$ , HR=3.84), Ann Arbor stage ( $P<0.05$ , HR=1.69), IPI score ( $P<0.05$ , HR=1.61), and type of ulcers ( $P<0.05$ , HR=1.34). It seems that the IPI score ( $P=0.018$ ) had a stronger effect on distinguishing patients with different prognoses than the Ann Arbor stage ( $P=0.045$ , marginally significant). Post-treatment complications seem to be more predictive of prognosis than complications before treatment: of the 14 patients with post-treatment complications, only 4 (28.6%) survived the 3-year threshold.

### Quantitative Data Processing

We used the pROC package of R 3.5.3 to draw ROC curves for some major quantitative data and obtained cut-off values. The mean age of patients was 61 years: 33 patients (40.2%) were younger than 60 and showed a 3-year survival rate of 72.7%, while 15 patients (18.3%) were over 75 and showed a 3-year survival rate of only 26.6%. The ROC curve of age for the 3-year survival rate suggests that the inflection point occurs at age 74.5, with an AUC of 0.691 (**Figure 1A**). The AUC of the ROC curve for Ki67 was only 0.548, suggesting that Ki67 had little effect on determining the prognosis. The median HGB level was 111 g/L, with a minimum value of 57 g/L and a maximum value of 167 g/L. The HGB-based ROC curve had an AUC of 0.684 and an inflection point at 104.5 g/L (**Figure 1B**).

The LDH values ranged from 125 to 1833 U/L with a median of 230.5 U/L and a mean of 346.8 U/L. The AUC of its ROC curve was 0.659 and the inflection point was at 301 U/L (**Figure 1C**). The median value of ALB was 36.9 g/L with a minimum value of 23.1 g/L and a maximum value of 46.5 g/L. The AUC of its ROC curve was 0.712 and the inflection point was at 38.15 g/L (**Figure 1D**). The median value of  $\beta 2$ MG was 2.22 mg/L with a minimum value of 0.94 mg/L, and a maximum value of 12.2 mg/L. The AUC of its ROC curve was 0.737 and the inflection point was 2.39 mg/L (**Figure 1E**). The median level of MIALD was 2.8 cm, with a maximum value of 17.5 cm. The AUC of the ROC curve was 0.638 and the inflection point was at 3.9 cm (**Figure 1F**). Based on these results, we selected the following cut-off values for the survival analysis: age 75 years, HGB level of 105 g/L, ALB level of 38 g/L,  $\beta 2$ MG level of 2.4 mg/L, and MIALD of 4 cm.

### Survival Analysis

Using the cut-off values selected, we were able to perform a univariate survival analysis on all variables and compare their *P* values. The significant variables (with *P* values less than 0.05) were: HGB level, age, LDH level, ALB level, MIALD, IPI score, primary gastrointestinal lymphoma (PGL; yes or no), Ann Arbor stage, and  $\beta 2$ MG level (**Figure 2A**). The effect of HGB was the most significant, followed by age and LDH level. Both the Ann Arbor stage and the IPI score were significant in assessing prognosis but were relatively weak, while the diagnosis of PGL (yes or no) was somewhere in between. However, the predictive value of ALB and MIALD for prognosis was better than the IPI score.

### The HLAMA and the Simplified Models

To identify the most significant variables for the prognostic evaluation, we established 2 models: the COX regression model and the random forest model. Based on the COX regression

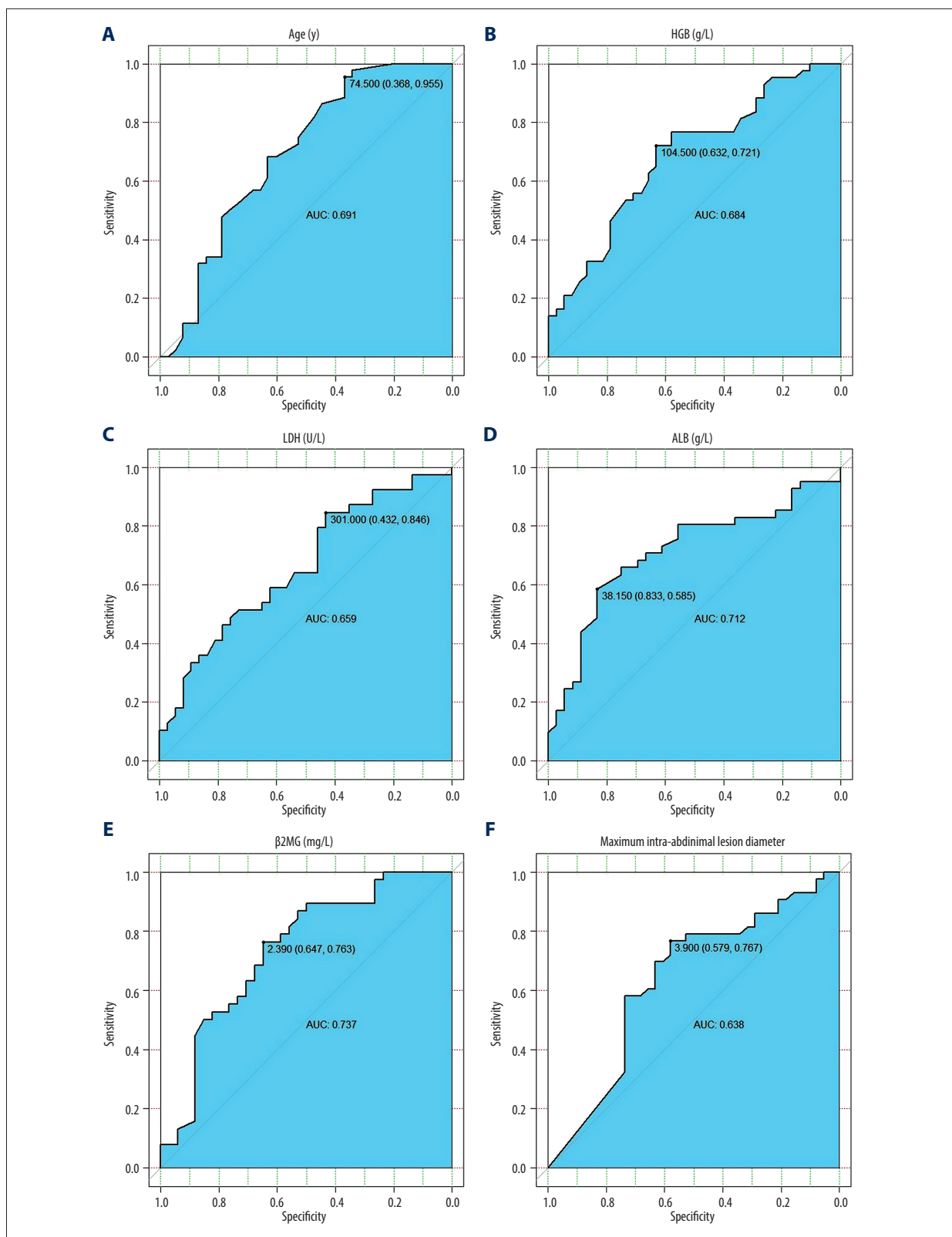
**Table 1.** Patient information and 3-year survival rate.

Clinical features	Number and proportion of patients		Number and proportion of patients that survived		$\chi^2$	p-value
Total	82		48 (58.5%)			
Age (years)						
≤60	33	(40.2%)	24	(72.7%)	9.0162	0.01102
Between 60-75	34	(41.5%)	20	(58.5%)		
>75	15	(18.3%)	4	(26.6%)		
Gender						
Male	45	(54.9%)	27	(60.0%)	0.0051	0.9431
Famale	37	(45.1%)	21	(56.8%)		
PGL						
Yes	38	(46.3%)	28	(73.6%)	5.5823	0.01814
No	44	(53.7%)	20	(45.5%)		
Region involved						
Stomach	42	(51.2%)	28	(66.7%)		0.2002
Small intestine	12	(14.6%)	4	(33.3%)		
Ileocecum	14	(17.1%)	8	(57.1%)		
Colon	4	(4.9%)	1	(25.0%)		
Rectum	1	(1.2%)	1	(100%)		
Multiple sites	9	(11.0%)	6	(66.7%)		
Symptoms						
Asymptomatic	6	(7.3%)	3	(50.0%)		0.7485
Non-specific symptoms	58	(70.7%)	33	(56.9%)		
Specific symptoms	18	(22.0%)	12	(66.7%)		
Pre-treatment complications						
Yes	33	(40.2%)	17	(51.5%)	0.68985	0.4062
Bleeding	20	(24.4%)	12	(60.0%)		
Perforation	0	(0.0%)	0	(0.0%)		
Obstruction	13	(15.9%)	4	(30.8%)		
Intussusception	5	(0.06%)	3	(60.0%)		
No	49	(59.8%)	31	(63.2%)		
Post-treatment complications						
Yes	14	(17.1%)	4	(28.6%)	4.8455	0.02772
No	68	(82.9%)	44	(64.7%)		

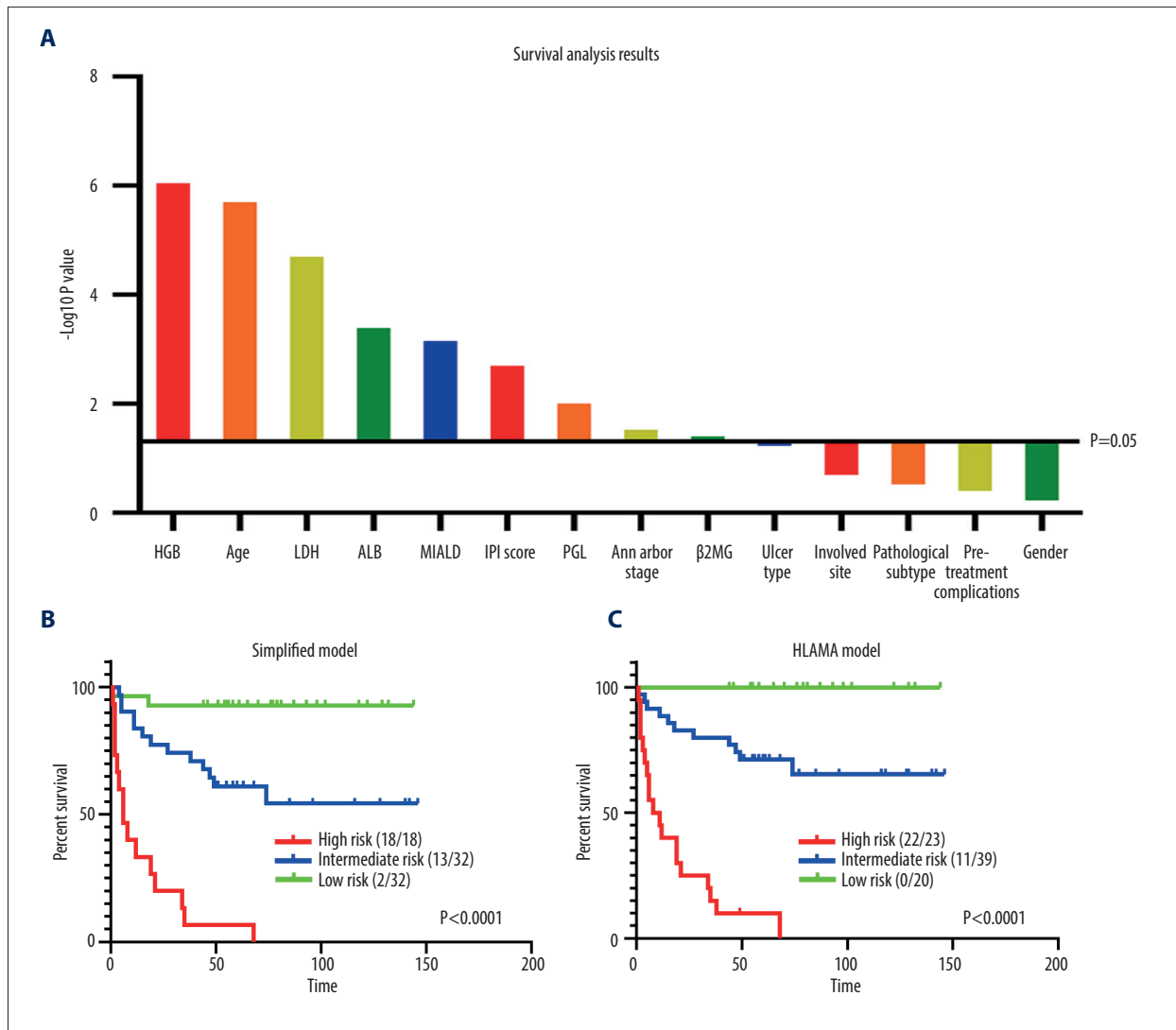
**Table 1 continued.** Patient information and 3-year survival rate.

Clinical features	Number and proportion of patients		Number and proportion of patients that survived		$\chi^2$	p-value
Pathological subtype						
ABC	38	(46.3%)	21	(55.3%)	0.34074	0.8434
GCB	24	(29.3%)	15	(62.5%)		
Not classified	20	(24.4%)	12	(60.0%)		
Bcl2						
Positive	30	(36.6%)	12	(40.0%)	2.9881	0.08388
Negative	22	(26.8%)	15	(68.2%)		
Bcl6						
Positive	34	(41.5%)	21	(61.8%)	0.53684	0.4637
Negative	34	(41.5%)	17	(50.0%)		
Ann Arbor stage						
I	17	(20.7%)	14	(82.4%)		0.04476
II	14	(17.1%)	10	(71.4%)		
III	9	(11.0%)	5	(55.6%)		
IV	42	(51.2%)	19	(45.2%)		
IPI score						
0	9	(11.1%)	8	(88.9%)		0.01874
1	13	(15.9%)	8	(61.5%)		
2	17	(20.7%)	13	(76.4%)		
3	15	(18.3%)	6	(40.0%)		
4	20	(24.4%)	8	(40.0%)		
5	2	(2.4%)	0	(0.0%)		
Types of ulcers						
No ulcer	22	(26.8%)	9	(40.9%)		0.02365
Simple ulcer	30	(36.6%)	24	(80.0%)		
Multiple ulcers	12	(14.6%)	6	(50.0%)		
Giant Ulcer	13	(15.9%)	7	(53.8%)		
Endoscopic findings						
Mucosal erythema/benign-appearing gastric ulcer	48	(58.5%)	31	(64.6%)	1.5733	0.4554
Thickened or diffuse bulge	15	(18.3%)	7	(46.7%)		
Mass	14	(%)	8	(57.1%)		

PGL – primary gastrointestinal lymphoma; specific symptoms include hematemesis, blood in the stool, melena, and cessation of exhaustion; ABC – activated B cell type; GCB – germinal center B cell type.



**Figure 1.** ROC curves of the variables used. **(A)** ROC curve of age and 3-year survival rate. **(B)** ROC curve of HGB and 3-year survival rate. **(C)** ROC curve of LDH and 3-year survival rate. **(D)** ROC curve of ALB and 3-year survival rate. **(E)** ROC curve of  $\beta$ 2MG and 3-year survival rate. **(F)** ROC curve of MIALD and 3-year survival rate.



**Figure 2.** Survival analysis. (A) Univariate analysis of all variables and their P values. (B) Survival analysis of the simplified model. (C) Survival analysis of HLAMA model.

model, HGB and LDH were selected as variables that could be included in the regression equation (Table 2,  $P < 0.01$ ). On the other hand, the random forest model, which was based on patient outcome and had numerical and categorical variables, had an accuracy rate in predicting the three-year survival rate of 76.5% and a recall rate of 81.3%. The most significant variables were HGB, MIALD, age, and ALB (Table 3, Figure 3).

Based on these results, we built 2 models that incorporated the most significant variables (HGB level, age, LDH level, ALB level, and MIALD): an integral model with 5 variables and a simplified model with only 3 variables. Given that the HGB regression coefficient was higher in the COX regression model (Table 2) and that it was the most significant variable in both models, we gave HGB the highest weight (weight=2), while the other variables were given lower weights (weight=1).

The model including all 5 variables was named HLAMA and in this model, HGB values  $< 105$  g/L scored 2 points; LDH levels  $\geq 300$  U/L, age  $\geq 75$  y, ALB  $< 38$  g/L, and a maximum intra-abdominal lesion diameter (MIALD)  $\geq 4$  cm scored 1 point each. The simplified model included only HGB level, LDH level, and age (Table 4). For the HLAMA model, a high-risk group scored 4 points or higher, while a score of 0 is considered a low-risk group and a score between 0 and 4 is considered as intermediate-risk. For the simplified model, the high-risk group scored 3 or higher, the low-risk group scored 0, and the intermediate-risk group scored between 0 and 3. Using these 2 models, we were able to distinguish patients with poor prognosis from patients with good prognosis. According to the HLAMA model, there were 23 high-risk patients, of whom 22 died and the median survival period was only 6 months. The median survival period for patients in the intermediate-risk group was

**Table 2.** Variables used in the COX regression.

		B	SE	Wald	Degree of freedom	P value	Exp(B)	95.0% Exp(B) CI	
								Up	Down
Step 1	HGB	-1.975	0.469	17.689	1	0	0.139	0.055	0.348
Step 2	HGB	-2.074	0.478	18.802	1	0	0.126	0.049	0.321
	LDH	-1.521	0.41	13.741	1	0	0.219	0.098	0.488

B indicates regression coefficient, SE indicates standard error, Exp(B) indicates relative risk.

**Table 3.** Variables in a random forest model.

	Alive	Dead	Mean decrease accuracy	Mean decrease gini
Age	4.2685533	2.0465212	4.3303862	4.165581
HGB	5.829504	9.4041157	9.3688256	5.993836
β2MG	2.4561263	3.1466944	3.8752958	3.659912
ALB	3.0848666	5.9015839	5.963592	4.29171
LDH	1.5605664	-1.7406013	0.1242688	2.424922
PGL	-0.2404816	-0.5040394	-0.4901279	0.592321
Pathological subtype	4.2577819	3.2987249	3.109322	1.24679
MIALD	6.1794461	3.5795777	5.002395	3.79236

Pathological subtype includes three groups: ABC, GCB, and not classified.

55 months. There were 20 low-risk patients and all who continued their follow-up survived (**Figure 2B, 2C**).

### Prognosis by Age Group and IPI Score

Since a larger age cut-off value is provided in the final model, we evaluated the effectiveness of the model in different age groups. Among the population with an age over 60 years, patients with good prognosis could still be selected by the HLAMA model (**Figure 4**). A similar distinction was found in the simplified model. On the other hand, both models included younger patients in the low- or intermediate-risk group.

We compared the high-, intermediate-, and low-risk groups of the model with the corresponding groups determined by IPI scores (**Figure 5A-5D**). The prognosis of the high-risk group based on the HLAMA model was significantly lower than the prognosis based on IPI scores ( $P=0.016$ ). The prognosis of the low-risk group based on the HLAMA model was significantly better than the prognosis based on the IPI score ( $P=0.04$ ). There was no significant difference in the prognosis of the intermediate-risk group between the 2 models. We also compared the groups obtained by NCCN-IPI scores with those obtained by HLAMA scores. Because the NCCN-IPI score divides patients into 5 risk stratifications, and extranodal involvement is used as an integral item, we compared the low-risk

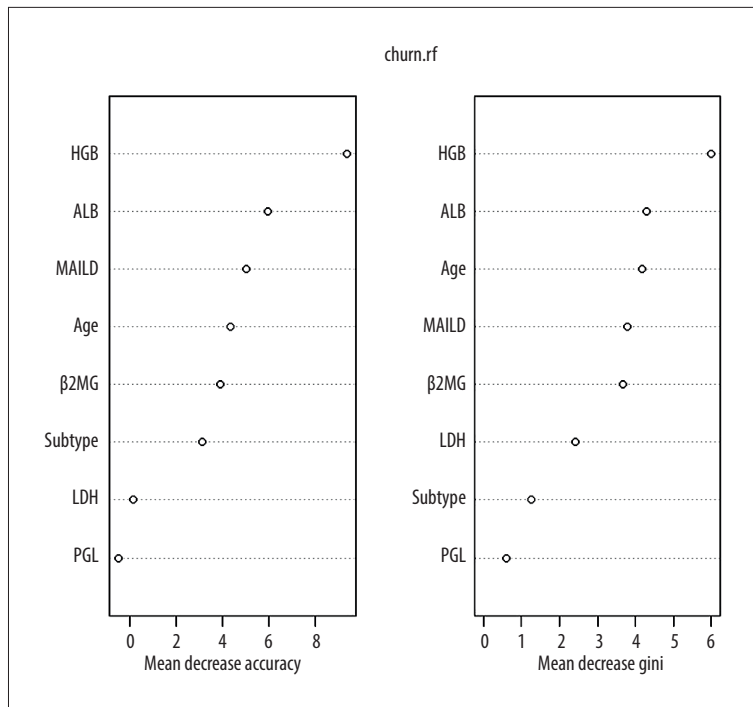
group of HLAMA with the medium-risk and low-risk groups of the NCCN-IPI score. The results showed that the survival of the HLAMA low-risk group was similar to that of the NCCN-IPI low-risk group, but more patients were enrolled. The HLAMA high-risk group had worse survival than the NCCN-IPI high-risk group ( $P=0.016$ ) (**Figure 5E, 5F**).

We are particularly concerned about the performance of this model in primary diffuse large B cell lymphoma of the gastrointestinal tract because the ability to identify relatively high-risk patients from PGL is one of its main clinical implications. The survival analysis based on different risk groups differed significantly in both the HLAMA model ( $p=0.0013$ ) and the simplified model ( $P<0.0001$ ) (**Figure 6**).

## Discussion

The gastrointestinal tract is the most common site of extranodal involvement in lymphoma cases, where the entire digestive tract can be involved. GI-DLBCL is an aggressive lymphoma that may be de novo or originate from another lymphoma (usually MALT lymphoma). It accounts for 40-70% of all gastric lymphomas and is more common in men ages 50-60 years [4,15]. DLBCL often shows significant heterogeneity, and diffuse large B cell lymphoma of the gastrointestinal tract is no





**Figure 3.** Variables in a random forest model. The 'Subtype' represents the pathological subtype.

**Table 4.** Variable values for the HLAMA model and for the simplified model.

	HLAMA score	Simplified model
HGB (g/L)	<105	2
	≥105	0
LDH (U/L)	≥300	1
	<300	0
Age	≥75y	1
	<75y	0
ALB (g/L)	<38	1
	≥38	0
MIALD (cm)	≥4	1
	<4	0

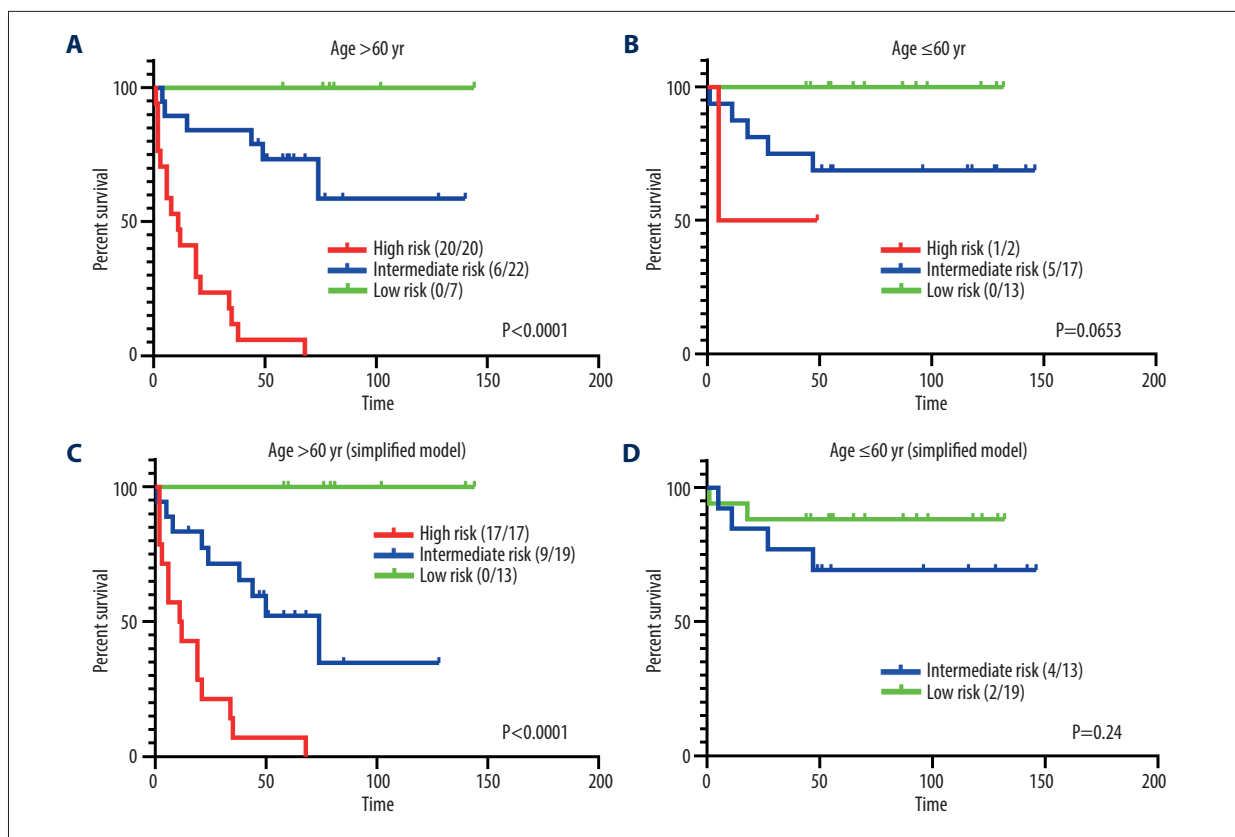
For the HLAMA model: ≥4 points indicate high-risk; 0 points indicate low-risk; 1 to 3 indicates intermediate-risk. For the simplified model: ≥3 points indicate high-risk; 0 points indicate low-risk; 1 or 2 indicates intermediate-risk.

exception [16]. Therefore, prognostic characterization of individual patients is necessary to determine the best treatment.

The prognosis of gastrointestinal lymphoma is evaluated by several classification methods, but the Ann Arbor staging system and the Musshoff-modified Ann Arbor staging system are still the most commonly used staging systems. In addition, the Lugano, the TNM, and the Paris staging systems are also

used, even though their accuracy is still controversial [4]. For instance, the prognostic ability of the TNM staging system may be weakened due to the surgical treatment of gastrointestinal lymphoma [17]; the Musshoff-modified Ann Arbor staging system may not provide sufficient prognostic indicators for early-stage patients with localized PG-DLBCL [7]; and, finally, the IPI score, which is widely used in the assessment of lymphoma prognosis, may not accurately distinguish the prognosis between high-risk patients and low- or intermediate-risk patients [18,19]. To solve these issues, Zhang et al [7] proposed a prognostic method that combines the IPI score with the Lugano staging system, but this made prognostic evaluation even more complicated [7]. Therefore, we aimed to design a more concise prognostic tool.

Several studies have evaluated the prognostic factors of gastrointestinal lymphoma using both univariate analysis and COX regression multivariate analysis [16,21-23]. In 2014, Song et al [22] conducted a study with 85 patients who had diffuse large B cell lymphoma of the gastrointestinal tract, and found that age, lesion site, tumor size, clinical Lugano staging of the gastrointestinal tract, and IPI score affected OS of patients. Additionally, Bautista-Quach et al [16] also found that age, stage of the disease, lactate dehydrogenase (LDH) levels, and chemotherapy treatment were independently and significantly related to survival [16]. Zhang et al [7] showed that age over 60 years, extranodal multi-site/gastrointestinal involvement, increased serum lactate dehydrogenase and β2-microglobulin, and decreased serum albumin were survival-related factors. In our study, we found that hemoglobin levels,



**Figure 4.** Prognostic analysis for different age groups. (A) HLAMA model in patients older than 60 years. (B) HLAMA model in patients younger than 60 years. (C) Simplified model in patients older than 60 years. (D) Simplified model in patients younger than 60 years.

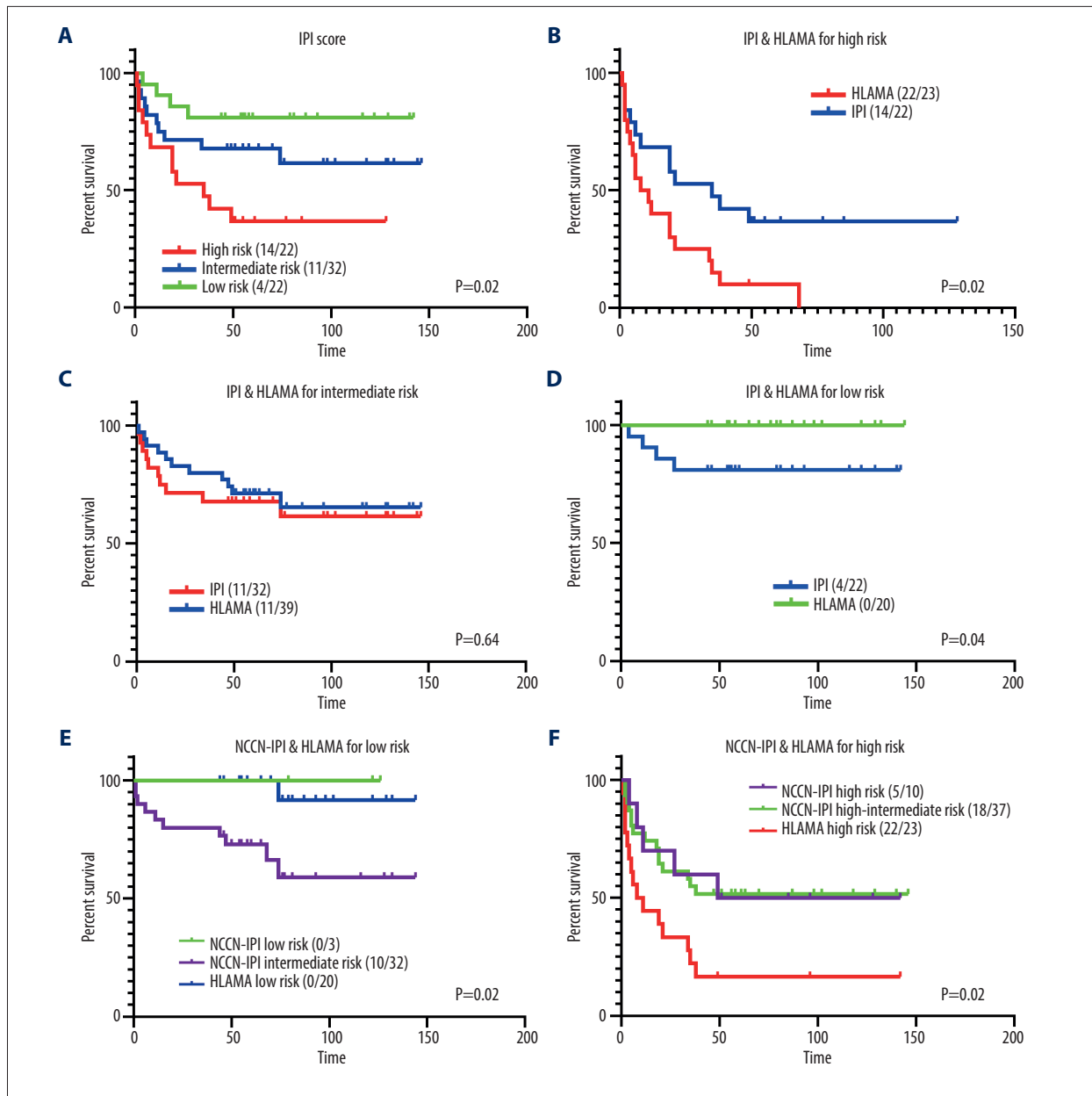
age, lactate dehydrogenase (LDH) levels, serum albumin levels, and the maximum intra-abdominal lesion diameter at initial diagnosis significantly affected the survival of patients. Hematologists may be familiar with these indicators, as hemoglobin levels and serum albumin levels have appeared in the IPS score of Hodgkin lymphoma [23], while age and LDH level are indicators of IPI score. Among the several prognostic models, the cut-off values of a given indicator are very similar, indicating the importance of these indicators.

The relationship between anemia and gastrointestinal lymphoma is not fully understood, but its manifestation is very common in several gastrointestinal tumors. Väyrynen et al [24] reported that 43% of patients with colorectal cancer developed anemia, the most common being normocytic anemia, followed by microcytic anemia. Their study suggests that decreased levels of hemoglobin are associated with systemic inflammation and that normocytic anemia is related to lower overall survival rates [24]. Patients with anemia do not always show significant manifestations of gastrointestinal bleeding. In fact, in this study, 39 patients had hemoglobin levels below 110 g/L, but only 15 (18%) had meaningful manifestations of gastrointestinal bleeding (including melena, vomiting, and positive

occult blood). Nonetheless, the relationship between anemia and gastrointestinal lymphoma needs further investigation.

Albumin level is frequently used in lymphoma and gastrointestinal tumor studies, and its importance was confirmed in a study with 75 cases of primary diffuse large B cell lymphoma of the stomach, in which the authors reported that ALB ≤35 g/L, stage ≥IIE, and multi-site involvement were factors predicting poor prognosis [25]. Coincidentally, Mao et al [26] also considered ALB as a predictor of poor prognosis for patients with primary intestinal diffuse large B cell lymphoma (PI-DLBCL). Similarly, our study suggests that ALB is only slightly reduced and may also indicate a poor prognosis (reference range, 40-55 g/L). Differences in the reference range may be due to differences in the detection methods used. Additionally, decreased albumin levels may indicate poor nutritional status or the loss of protein caused by tumors, or may be related to the risk of thrombosis [27]. Nonetheless, decreased levels of albumin might indicate a greater inflammatory response, potentially leading to poorer outcomes [28].

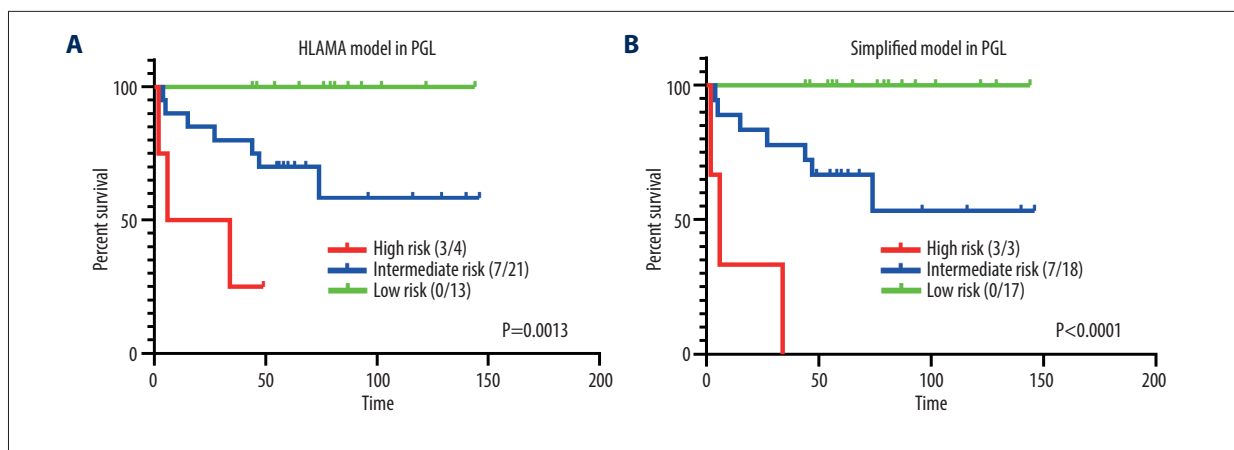
Inclusion of the largest measurable lesion diameter (MIALD) into the model was based on the widespread application of



**Figure 5.** Comparison with the IPI score. **(A)** Risk stratification based on IPI score. **(B)** Comparison of high-risk group obtained with the HLAMA score and with the IPI score. **(C)** Comparison of intermediate-risk group obtained with the HLAMA score and with the IPI score. **(D)** Comparison of the low-risk group obtained with the HLAMA score and with the IPI score. **(E)** Comparison of low-risk group obtained with the HLAMA score and with the NCCN-IPI score. **(F)** Comparison of the high-risk group obtained with the HLAMA score and with the NCCN-IPI score.

the Lugano staging system in recent years. In addition, including such a simple and easy-to-obtain image result is very useful for clinical applications. A recent study has demonstrated that prognosis can be assessed in lymphoma clinical trials using the sum of the largest diameters of up to 3 target lesions [29]. Our study also found that patients with MIALD greater than 4 cm had worse prognosis. In recent years, imaging indicators, such as the maximum SUV value and lesion diameter, have often been used.

We developed a 5-variable HLAMA model and a 3-variable simplified model: the HLAMA model basically fits the results of the random forest model, while the simplified model basically fits the COX regression model. Setting up a more complex HLAMA model or adding age to the simplified model was not carried out in order to avoid overfitting caused by too few variables and to reduce the bias of the data itself. We expect other researchers to test and modify our model to create a more accurate and concise model.



**Figure 6.** HLAMA model and simplified model of patients with primary gastrointestinal lymphoma. **(A)** Comparison of different risk groups with PGL according to the HLAMA model. **(B)** Comparison of different risk groups with PGL according to the simplified model.

Regardless of the model used, we recommend that patients be distinguished in terms of risk groups. The median survival time of patients in the high-risk group, according to the HLAMA model, was only 6 months and most patients did not survive. In addition to effectively screening high-risk patients, the model also shows significant differences between low- and intermediate-risk patients. Of the 39 patients in the intermediate-risk group, nearly half belonged to stage I-II of the Ann Arbor staging system, and 8 patients belonged to the low-risk group based on the IPI score. Thus, distinguishing such groups of patients is the best practice.

One potential flaw of this study is the age of our patients, as the mean age was 61 years and, in previous studies, the mean age of gastrointestinal lymphoma patients was around 55 [30]. However, we found that patients with gastrointestinal lymphoma can tolerate treatment and achieve better prognosis even in this age-specific group: of the 49 patients over age 60 years, 29 were still in the low-risk group. Recently, a new prognosis score, the International Prognostic Index of the National Comprehensive Cancer Network (NCCN-IPI), was given to 1660 DLBCL patients treated with R-CHOP at the British Columbia Institute. Based on the linear effect of age on survival, NCCN-IPI divides advanced age into different categories (40-60 years, 60-75 years, and over 75 years), which are associated with increased risk [31]. In our study, the NCCN-IPI score was compared with the HLAMA score. Although NCCN-IPI has carried out 5 stratifications, the effect is relatively poor in GI-DLBCL patients, and the identification of high-risk patients is not as good as the HLAMA score, and because all patients have extranodal involvement, the number of patients included in the low-risk group with NCCN-IPI score is insufficient.

In addition, the number of patients included in this study was relatively small, and data related to molecular markers and/or tumor immune microenvironment were not included, which may limit the application potential of the scoring system. We aim to make up for these shortcomings in prospective follow-up studies, and look forward to multi-center cooperation.

## Conclusions

We developed a prognostic model for diffuse large B cell lymphoma of the gastrointestinal tract based on hemoglobin levels, age, lactate dehydrogenase (LDH) levels, serum albumin levels, and the maximum intra-abdominal lesion diameter. We believe that both the HLAMA model and its simplified version are practical for clinical use. Both models are similar to the IPI score, but could be considered better as they can provide a simpler and more accurate prognostic assessment in patients with gastrointestinal lymphoma. For patients with localized gastrointestinal lymphoma, our model can identify high-risk patients, and for relatively older patients, it can identify the low-risk population, enabling the possibility of more aggressive treatment.

## Acknowledgments

We would like to thank Mr. Kai Hu from Beijing Boren Hospital, for support to this work.

## Conflicts of Interest

None.

## References:

1. Castillo JJ, Winer ES, Olszewski AJ. Sites of extranodal involvement are prognostic in patients with diffuse large B-cell lymphoma in the rituximab era: An analysis of the Surveillance, Epidemiology and End Results database. *Am J Hematol*. 2014;89(3):310-14
2. Carbone PP, Kaplan HS, Musshoff K, et al. Report of the committee on Hodgkin's disease staging classification. *Cancer Res*. 1971;31:1860-61
3. Musshoff K, Schmidt-Vollmer H. Proceedings: Prognosis of non-Hodgkin's lymphomas with special emphasis on the staging classification. *Z Krebsforsch Klin Onkol Cancer Res Clin Oncol*. 1975;83(4):323-41
4. Psyrri A, Papageorgiou S, Economopoulos T. Primary extranodal lymphomas of stomach: Clinical presentation, diagnostic pitfalls and management. *Ann Oncol*. 2008;19:1992-99
5. Rohatiner A, d'Amore F, Coiffier B, et al. Report on a workshop convened to discuss the pathological and staging classification of gastrointestinal tract lymphoma. *Ann Oncol*. 1994;5:397
6. Chang S, Shi X, Xu Z, et al. TNM staging system may be superior to Lugano and Ann Arbor systems in predicting the overall survival of patients with primary gastrointestinal lymphoma. *J BUON*. 2015;20(3):812-19
7. Zhang S, Wang L, Yu D, et al. Localized primary gastrointestinal diffuse large B cell lymphoma received a surgical approach: An analysis of prognostic factors and comparison of staging systems in 101 patients from a single institution. *World J Surg Oncol*. 2015;13:246
8. Lossos IS, Morgensztern D. Prognostic biomarkers in diffuse large B-cell lymphoma. *J Clin Oncol*. 2006;24(6):995
9. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med*. 1993;329(14):987
10. Dawson IM, Cornes JS, Morson BC. Primary malignant lymphoid tumours of the intestinal tract. Report of 37 cases with a study of factors influencing prognosis. *Br J Surg*. 1961;49:80
11. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103:275
12. Meyer PN, Fu K, Greiner TC, et al. Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B cell lymphoma treated with rituximab. *J Clin Oncol*. 2011;29:200
13. Robin X, Turck N, Hainard A, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77
14. Breiman, L. Random forests. *Machine Learning*. 2001;45(1):5-32
15. Ferreri AJ, Montalbán C. Primary diffuse large B-cell lymphoma of the stomach. *Crit Rev Oncol Hematol*. 2007;63:65-71
16. Bautista-Quach MA, Ake CD, Chen M, et al. Gastrointestinal lymphomas: Morphology, immunophenotype and molecular features. *J Gastrointest Oncol*. 2012;3(3):209-25
17. Avilés A, Nambo MJ, Neri N, et al. The role of surgery in primary gastric lymphoma: Results of a controlled clinical trial. *Ann Surg*. 2004;240(1):44
18. Tibiletti MG, Martin V, Bernasconi B, et al. BCL2, BCL6, MYC, MALT1, and BCL10 rearrangements in nodal diffuse large B-cell lymphomas: A multi-center evaluation of a new set of fluorescent in situ hybridization probes and correlation with clinical outcome. *Hum Pathol*. 2009;40:645-52
19. Tay K, Tai D, Tao M, et al. Relevance of the International Prognostic Index in the rituximab era. *J Clin Oncol*. 2011;29:e14; author reply e15
20. Ziepert M, Hasenclever D, Kuhnt E, et al. Standard International Prognostic Index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. *J Clin Oncol*. 2010;28:2373-80
21. Magnoli F, Bernasconi B, Vivian L, et al. Primary extranodal diffuse large B-cell lymphomas: Many sites, many entities? Clinico-pathological, immunohistochemical and cytogenetic study of 106 cases. *Cancer Genet*. 2018;228-29:28-40
22. Song L, Cen X, Ou J, et al. [Long term follow-up and prognostic analysis of 85 cases with primary gastrointestinal diffuse large B cell lymphoma.] *Zhonghua Xue Ye Xue Za Zhi*. 2014;35(10):909-13 [in Chinese]
23. Hasenclever D, Diehl V. A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. *N Engl J Med*. 1998;339:1506
24. Väyrynen JP, Tuomisto A, Väyrynen SA, et al. Preoperative anemia in colorectal cancer: Relationships with tumor characteristics, systemic inflammation, and survival. *Sci Rep*. 2018;8(1):1126
25. Liu YZ, Xue K, Wang BS, et al. The size and depth of lesions measured by endoscopic ultrasonography are novel prognostic factors of primary gastric diffuse large B-cell lymphoma. *Leuk Lymphoma*. 2019;60(4):934-39
26. Mao L, Wang X, Wang CY, et al. [Evaluation of different staging systems and prognostic analysis of 110 primary gastrointestinal diffuse large B cell lymphoma.] *Zhonghua Yi Xue Za Zhi*. 2019;99(24): 1853-58 [in Chinese]
27. Takayoshi K, Kusaba H, Aikawa T. Hypoalbuminemia for the prediction of venous thromboembolism and treatment of direct oral anticoagulants in metastatic gastric cancer patients. *Gastric Cancer*. 2019;22(5):988-98
28. Kim S, McClave SA, Martindale RG. Hypoalbuminemia and clinical outcomes: What is the mechanism behind the relationship? *Am Surg*. 2017;83(11):1220-27
29. Younes A, Hilden P, Coiffier B, et al. International Working Group consensus response evaluation criteria in lymphoma (RECIL 2017). *Ann Oncol*. 2017;28(7):1436-47
30. Ding W, Zhao S, Wang J, et al. Gastrointestinal lymphoma in southwest China: Subtype distribution of 1,010 cases using the WHO (2008) Classification in a single institution. *Acta Haematol*. 2016;135(1):21-28
31. Zhou Z, Sehn LH, Rademaker AW, et al. An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood*. 2014;2014(123):837-42