Supplementary Materials

I. Supplementary Methods

II. Supplementary Reference

III. Supplementary Figures

IV. Supplementary Tables

V. Model use in clinical practice

Supplementary Methods

Selection of the Form of a Blood Parameter (Continuous or Dichotomous):

Although the use of continuous variables for further analysis can increase the flexibility of the model, the complexity of the model will increase accordingly. Akaike Information Criterion (AIC) can take into account both the model's flexibility and complexity. By minimizing the AIC values, the best form for each blood parameter can be determined.

Selection of Significant Blood Parameters and Model Construction:

We applied the Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression model in the developing data to select the significant variables with non-zero coefficients that can accurately predict the prognosis of GC[1-3]. The method uses an L1 penalty to shrink some regression coefficients to exactly zero. The penalty parameter λ, called the tuning parameter, controls the amount of shrinkage. With larger λ, the estimates of weaker factors shrink towards zero, so that only the strongest predictors remain in the model. The optimal values of the penalty parameter λ were determined by tenfold cross-validations. We selected λ via 1-SE (standard error) criteria, i.e., the optimal λ is the largest value for which the partial likelihood deviance is within one SE of the smallest value of partial likelihood deviance. A value λ = 0.086 with log (λ) = -2.454 was chosen by cross-validation via the 1-SE criteria. A vertical line was drawn at log (λ) = -2.454, which corresponds to the optimal value λ = 0.086 (Fig.1). The optimal tuning parameter resulted in five non-zero coefficients. Five blood parameters, Albumin, LMR, NLR, CEA, and CA19-9 with coefficients -0.86259331, 0.08074780, 0.06002645, 0.04007563, and 0.05939068, respectively, were selected in the LASSO Cox regression model.

Supplementary Reference:

[1] Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997; 16(4):385-95.

[2] Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2011; 73:273-282.

[3] Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. Biom J 2010; 52: 70–84.

Supplementary Tables

Supplementary table 1. Continuous forms of 9 blood parameters.

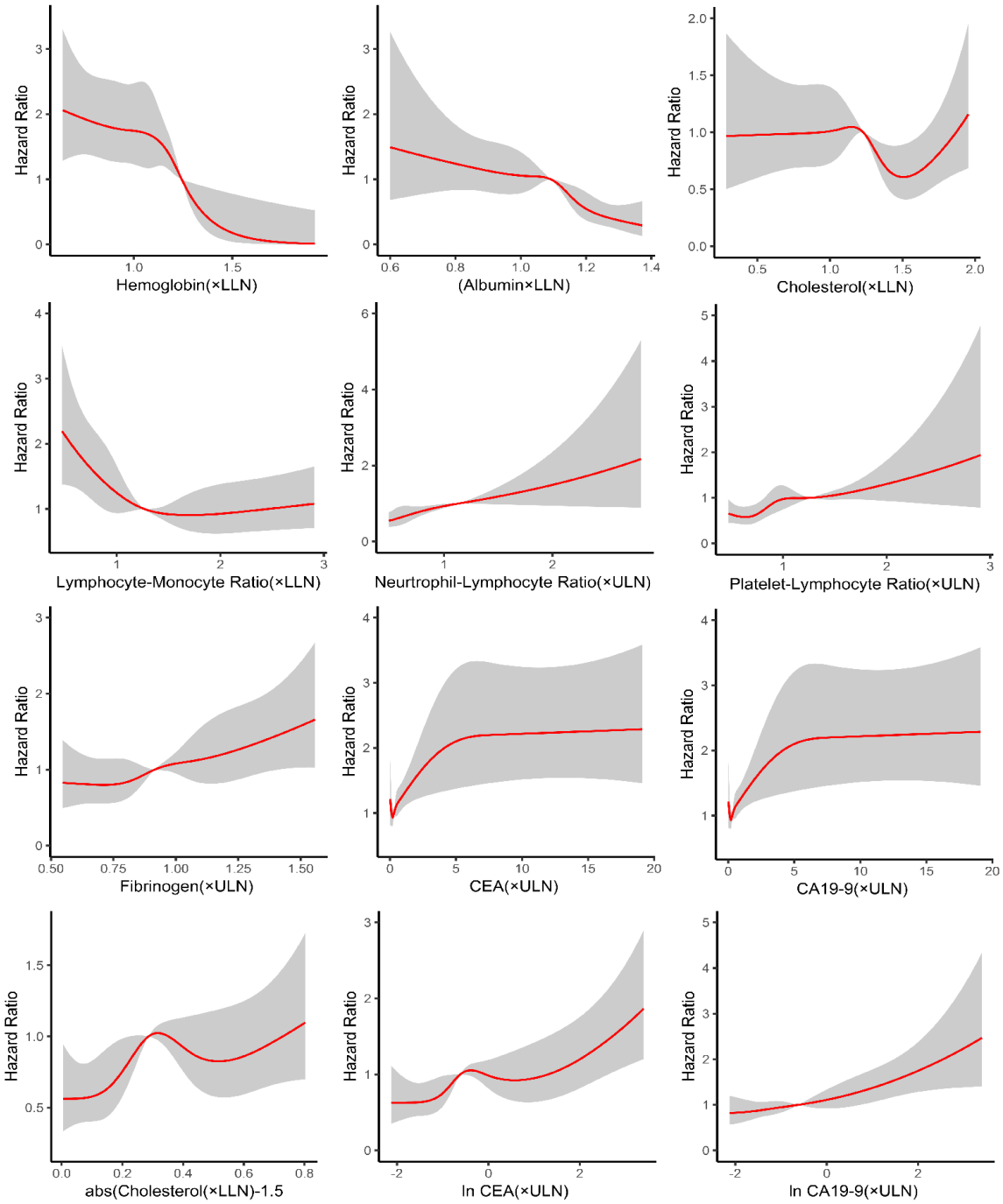
Variables	Coding (x stands for original value)	Additional transformation
Hemoglobin (g/L)	= x/120(110) (reference value(LLN):120 in men or 110 in women)	
Albumin (g/L)	= x/35 (reference value(LLN):35)	
Cholesterol (mmol/L)	= x/3.9 (reference value(LLN):3.9)	= abs(x/3.9-1.5)
LMR	= x/3.2 (cut-off value:3.2)	
NLR	= x/3.9 (cut-off value:3.9)	
PLR	= x/161 (cut-off value:161)	
Fibrinogen (g/L)	= x/4 (reference value(ULN):4)	
CEA (ng/mL)	= x/5 (reference value(ULN):5)	= ln x/5
CA19-9 (U/mL)	= x/37 (reference value(ULN):37)	= ln x/37

Supplementary table 2. The comparison of the Akaike Information Criterion (AIC) score and Harrell's c-statistic between the dichotomous and continuous forms for 9 blood parameters. A P-value < 0.05 indicates that Harrell's c-statistic of the continuous form of a blood parameter is significantly higher than that of the dichotomous form.

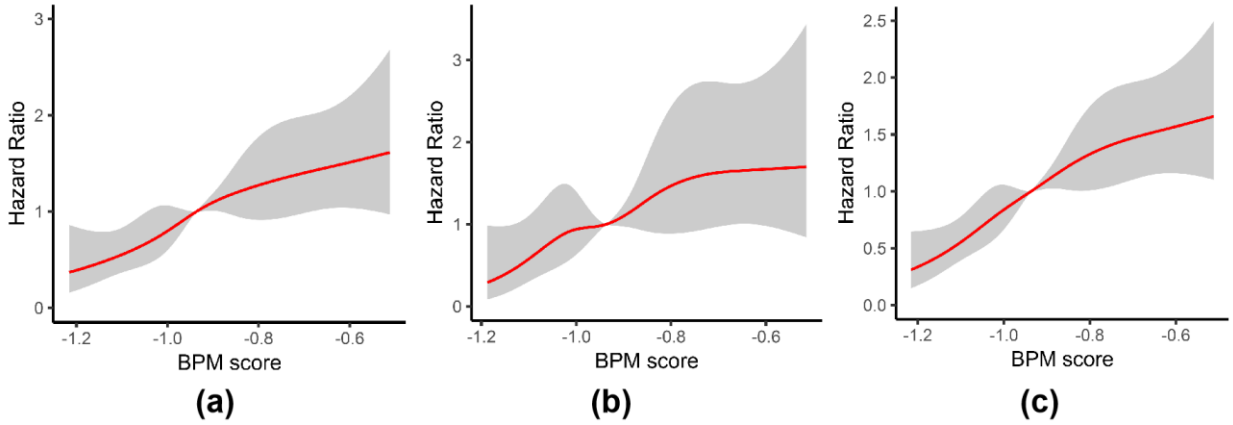
Variables	AIC		Harrel's C-statistic		P-value
	Dichotomic	Continous	Dichotomic	Continous	
Hemoglobin	3230.9	3225.3	0.549	0.586	**0.000**
Albumin	3224.4	3212.8	0.554	0.608	**0.000**
Cholesterol	3232.8	3235.0	0.539	0.552	0.176
LMR	3220.4	3233.3	0.564	0.565	0.896
NLR	3224.7	3230.9	0.535	0.547	0.312
PLR	3225.1	3225.9	0.558	0.572	0.209
Fibrinogen	3234.8	3230.0	0.529	0.552	**0.026**
CEA	3235.1	3224.8	0.533	0.569	**0.001**
CA19-9	3225.7	3230.7	0.550	0.552	0.826

Supplementary Figures

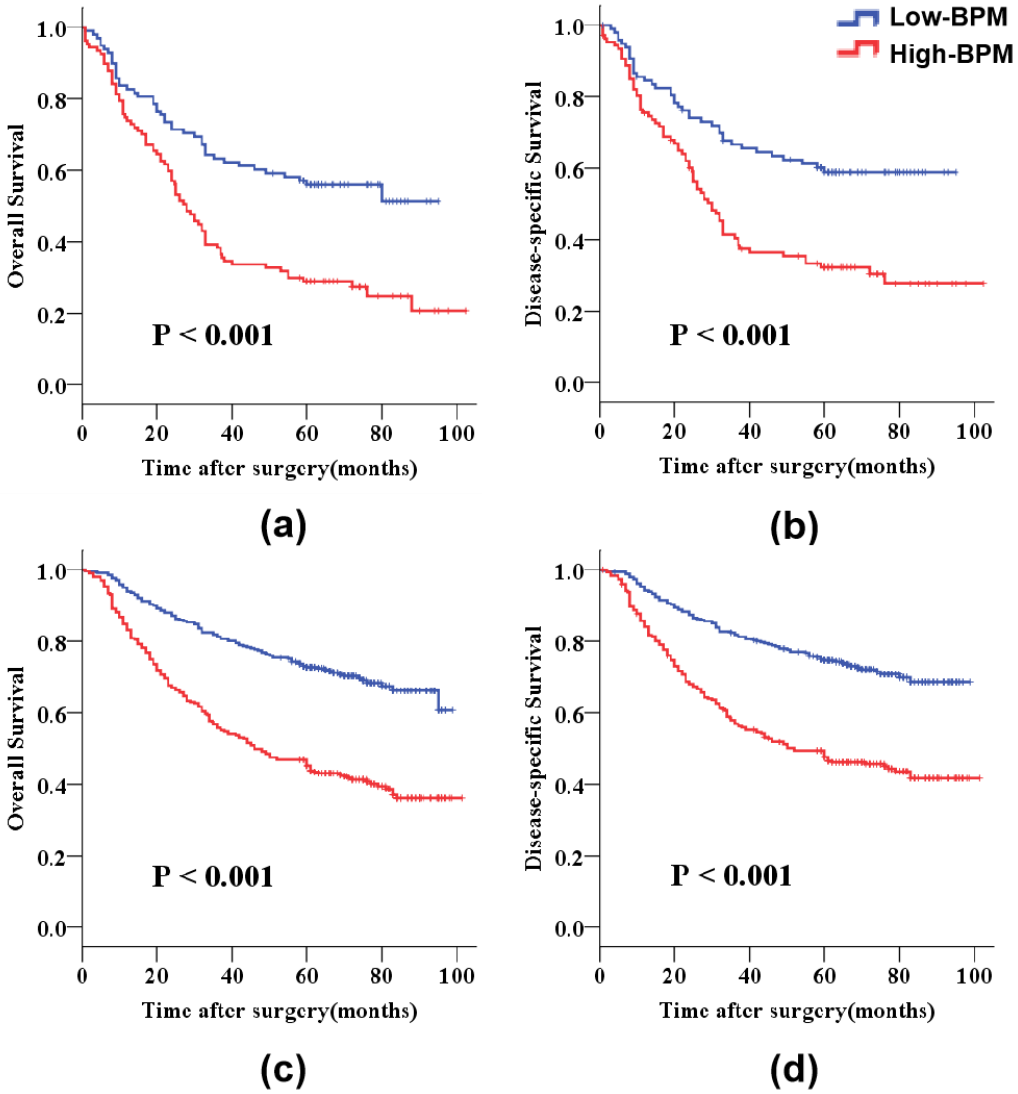
Supplementary figure 1. Univariate Cox analysis of OS with RCS in the developing cohort. After transforming, no strong non-linear effect was observed in all blood parameters.



Supplementary figure 2. Multivariate Cox analysis of OS with RCS in the developing (a), validation (b), and entire cohorts (c).



Supplementary figure 3. Kaplan-Meier survival analysis of overall survival and disease-specific survival in Adjuvant-Chemotherapy (AC) and non-AC groups. (**a**+**b**: AC group; **c**+**d**: non-AC group)



Model use in clinical practice

The following example explains how the BPM score can be used in clinical practice:

Suppose a 54-year-old male patient visits the outpatient clinic. He was recently diagnosed with GC, and his blood tests show the following results:

- Albumin (g/L) : 38.8 (reference value (LLN): 35)
- LMR : 2.32 (cutoff value: 3.2)
- NLR : 4.45 (cutoff value: 3.9)
- CEA (ng/mL) : 8.9 (reference value (LLN): 5)
- CA19-9 (U/mL) : 18.41 (reference value (LLN): 5)

Then we transform this 5 blood parameters according to the methods described in the article:

- Albumin (g/L) : = 38.8/35 = 1.11
- LMR : 1 (< 3.2)
- NLR : 1 (> 3.9)
- CEA (ng/mL) : = ln (8.9/5) = 0.58 (reference value (LLN): 5)
- CA19-9 (U/mL) : 0 (< 37)

We get the above numbers to be entered into the formula of BPM score:

- BPM = -0.86259331×1.11 + 0.08074780×1 + 0.06002645×1 + 0.05939068×0 + 0.04007563×0.58 = -0.793 (> -0.93)

This patient is classified as high-BPM group. After a planned curative-intent surgery, he is pathologically diagnosed as satge IIIc disease. To calculate the probability to survive for 1, 3 and 5 years, we evaluate the nomogram that combined the BPM score and TNM stage:

- BPM score : 5 (-0.793)
- TNM stage : 8 (IIIc)

The final score for the nomogram is 13, equal to a 70% survival rate for 1 year, a 30% survival rate for 3 years, and a 20% survival rate for 5 years.