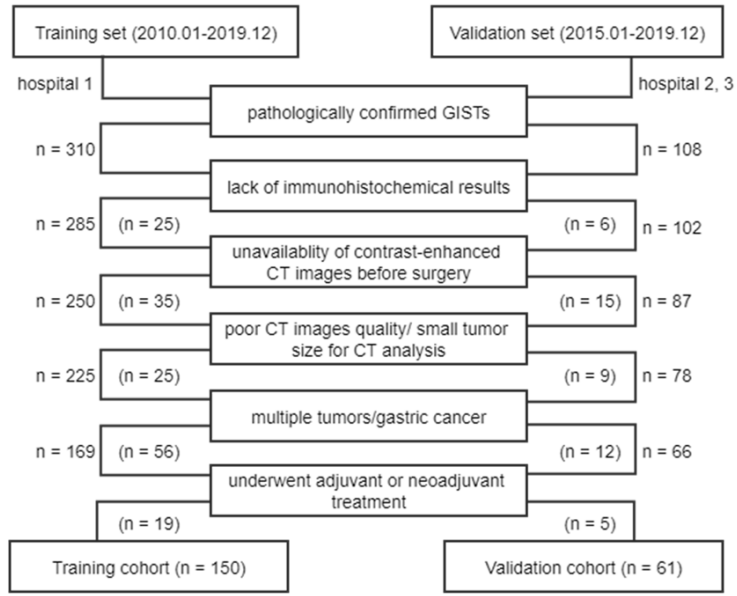


# A practical multi-class scoring system of gastric GISTs



hospital 1: Second Affiliated Hospital of Zhejiang University Medical School  
 hospital 2: Second Affiliated Hospital of Zhejiang Chinese Medical University  
 hospital 3: Hubei University of Medicine affiliated Renmin Hospital

Scheme S1. Workflow chart of patient selection and the exclusion criteria.

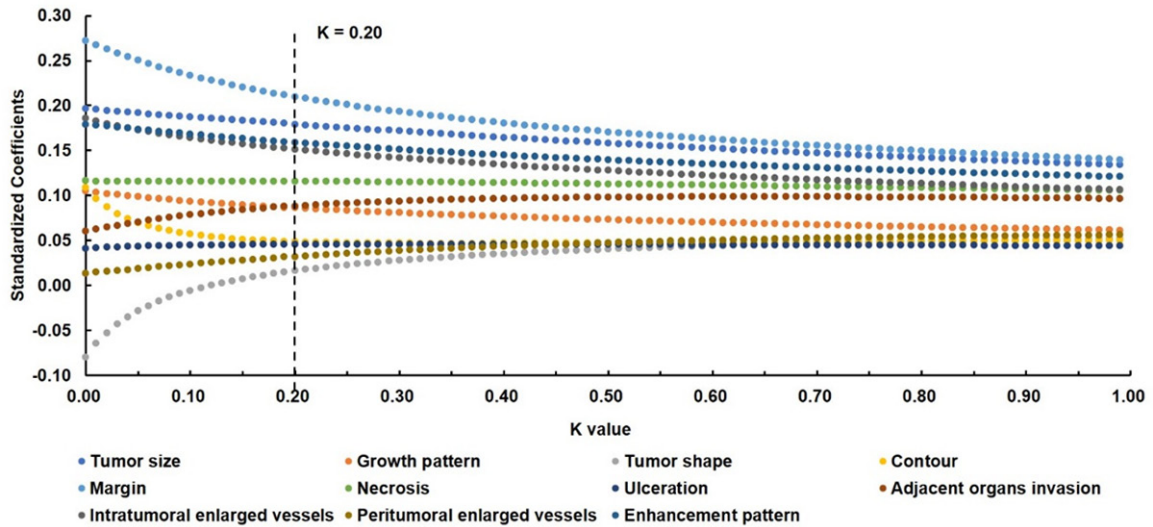
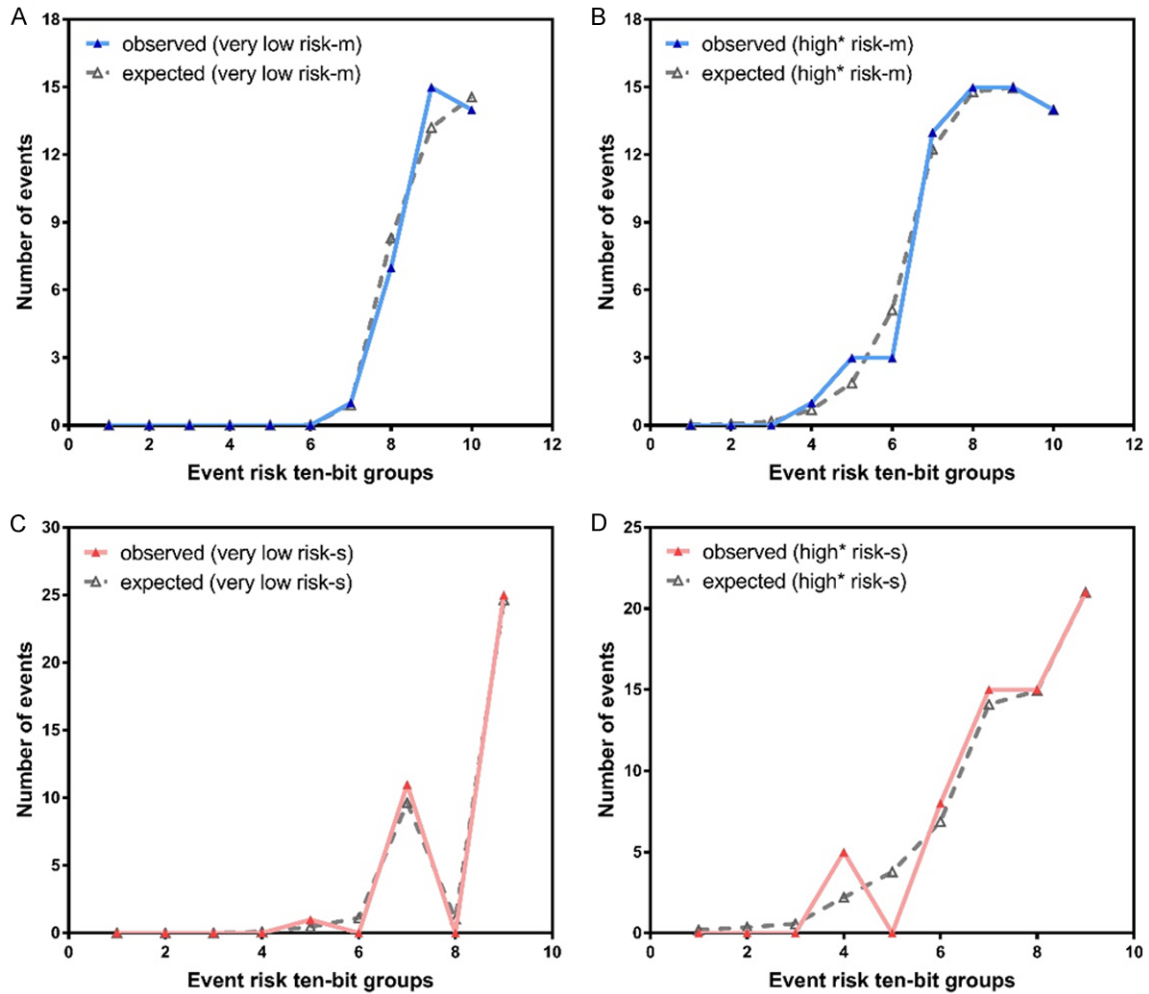


Figure S1. The ridge curve of the relevant predictors in gastric GISTs.

## A practical multi-class scoring system of gastric GISTs

**Table S1.** Ridge regression results of CT features (k = 0.20)

|                               | Unstandardized Coefficients |       | Standardized Coefficients | t     | p       | R <sup>2</sup> | Adjusted R <sup>2</sup> | p         |
|-------------------------------|-----------------------------|-------|---------------------------|-------|---------|----------------|-------------------------|-----------|
|                               | B                           | S.E.  | Beta                      |       |         |                |                         |           |
| Tumor Size                    | 0.034                       | 0.009 | 0.183                     | 3.748 | < 0.001 | 0.660          | 0.633                   | p < 0.001 |
| Growth pattern                | 0.132                       | 0.066 | 0.089                     | 1.991 | 0.048   |                |                         |           |
| Tumor shape                   | 0.019                       | 0.13  | 0.007                     | 0.143 | 0.886   |                |                         |           |
| Contour                       | 0.134                       | 0.131 | 0.051                     | 1.027 | 0.306   |                |                         |           |
| Margin                        | 0.556                       | 0.136 | 0.22                      | 4.086 | < 0.001 |                |                         |           |
| Necrosis                      | 0.258                       | 0.122 | 0.116                     | 2.114 | 0.036   |                |                         |           |
| Ulceration                    | 0.117                       | 0.123 | 0.045                     | 0.952 | 0.343   |                |                         |           |
| Adjacent organs invasion      | 0.209                       | 0.131 | 0.084                     | 1.592 | 0.114   |                |                         |           |
| Intratumoral enlarged vessels | 0.449                       | 0.134 | 0.157                     | 3.338 | 0.001   |                |                         |           |
| Peritumoral enlarged vessels  | 0.065                       | 0.117 | 0.028                     | 0.558 | 0.578   |                |                         |           |
| Enhancement pattern           | 0.358                       | 0.115 | 0.163                     | 3.103 | 0.002   |                |                         |           |
| Constant                      | 0.478                       | 0.09  | -                         | 5.323 | < 0.001 |                |                         |           |



**Figure S2.** A. The calibration curve of predictive model (m) in very low risk grade (p = 0.920). B. The calibration curve of predictive model (m) in high\* risk grade (p = 0.936). C. The calibration curve of score model (s) in very low risk grade (p = 0.721). D. The calibration curve of score model (s) in high\* risk grade (p = 0.098).

## A practical multi-class scoring system of gastric GISTs

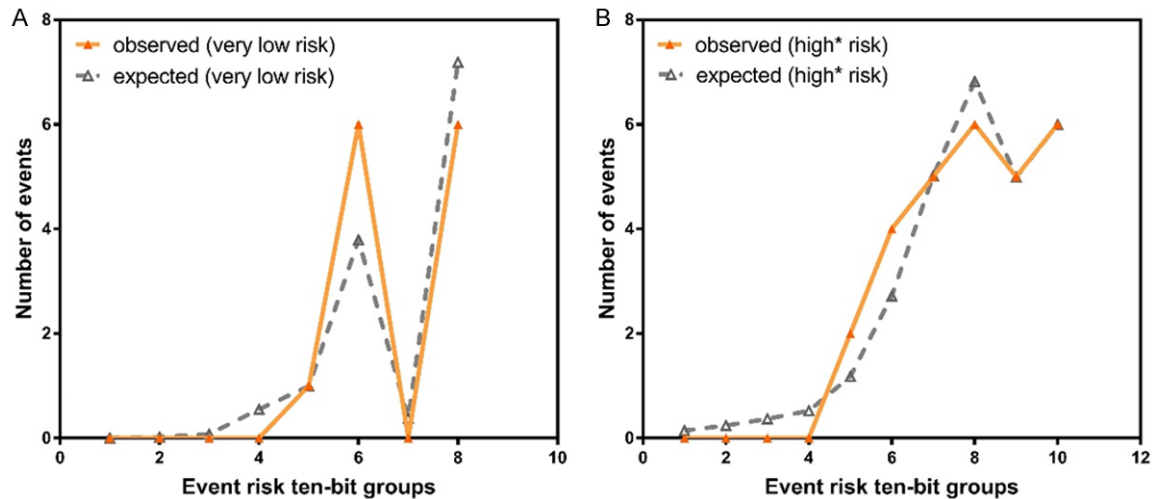
**Table S2.** The AUCs, sensitivity and specificity of predictive models and score models in training and validation cohorts

| Models                                   | AUC   | 95% C.I. |       | cut off point |             |
|--|-------|----------|-------|---------------|-------------|
|  |       | Lower    | Upper | sensitivity   | specificity |
| very low risk model for training         | 0.986 | 0.952    | 0.998 | 94.59%        | 98.23%      |
| very low risk score model for training   | 0.973 | 0.932    | 0.992 | 97.30%        | 93.81%      |
| high1* risk model for training           | 0.976 | 0.937    | 0.994 | 89.06%        | 97.67%      |
| high1* risk score model for training     | 0.977 | 0.938    | 0.994 | 92.19%        | 94.19%      |
| very low risk score model for validation | 0.912 | 0.833    | 0.982 | 92.31%        | 85.42%      |
| high1* risk score model for validation   | 0.972 | 0.894    | 0.997 | 100.00%       | 87.88%      |

**Table S3.** The predicted positive rates including precision, recall and F1 score in three score ranges of the training cohort

| Score range           | Predicted true positive | Actual positive | Precision | Predicted total positive | Recall | F1 score |
|-----------------------|-------------------------|-----------------|-----------|--------------------------|--------|----------|
| $\geq 0$ and $\leq 3$ | 36 <sup>1</sup>         | 37 <sup>1</sup> | 97.3%     | 43 <sup>1,2</sup>        | 83.7%  | 0.900    |
| $> 3$ and $\leq 8$    | 37 <sup>2</sup>         | 49 <sup>2</sup> | 75.5%     | 48 <sup>1,2,3</sup>      | 77.1%  | 0.763    |
| $> 8$ and $\leq 21$   | 54 <sup>3</sup>         | 64 <sup>3</sup> | 84.4%     | 59 <sup>2,3</sup>        | 91.5%  | 0.878    |

1: very low risk; 2: low risk; 3: high\* risk.



**Figure S3.** A. The calibration curve of score model in very low risk grade ( $p = 0.743$ ). B. The calibration curve of score model in high\* risk grade ( $p = 0.533$ ).

**Table S4.** The predicted positive rates including precision, recall and F1 score in three score ranges of the validation cohort

| Score range           | Predicted true positive | Actual positive | Precision | Predicted total positive | Recall | F1 score |
|-----------------------|-------------------------|-----------------|-----------|--------------------------|--------|----------|
| $\geq 0$ and $\leq 3$ | 12 <sup>1</sup>         | 13 <sup>1</sup> | 92.3%     | 19 <sup>1,2</sup>        | 63.2%  | 0.740    |
| $> 3$ and $\leq 8$    | 11 <sup>2</sup>         | 20 <sup>2</sup> | 55.0%     | 16 <sup>1,2,3</sup>      | 68.8%  | 0.611    |
| $> 8$ and $\leq 21$   | 24 <sup>3</sup>         | 28 <sup>3</sup> | 85.7%     | 26 <sup>2,3</sup>        | 92.3%  | 0.907    |

1: very low risk; 2: low risk; 3: high\* risk.