1. Supplementary Methods

CT acquisition and gastric preparation

Patients fasted for at least 6 h before CT examination. Unless contraindicated, 10 mg of racemic anisodamine hydrochloride was administered intramuscularly to reduce gastrointestinal peristalsis. Patients then received 6 g of effervescent granules with 10 mL of water for gastric distension. Adequate gastric distension is essential for reliable evaluation of the gastric wall, because a collapsed or underdistended stomach can mimic or obscure mural thickening and compromise the reproducibility of tumour measurements. Effervescent granules provide rapid and well-tolerated distension, with the released gas serving as negative intraluminal contrast against the enhancing gastric wall. CT examinations were performed from the diaphragmatic dome to the inferior pubic symphysis using GE LightSpeed 64 VCT or GE Discovery CT750 HD scanners. Nonionic iodinated contrast material was administered intravenously at 1.5 mL/kg and a rate of 3.5 mL/s. Arterial- and venous-phase images were acquired 35–40 s and 70 s after contrast injection, respectively. Multiplanar reconstruction was performed using a slice thickness of 5 mm.

Method S1. Candidate predictor selection and model specification

Candidate predictors were data-informed. Univariable association with CLDN18.2 positivity at $P<0.05$ was considered together with clinical relevance, availability at the intended time of model use, measurement reliability, redundancy with related variables, and model complexity relative to the number of outcome events. Multicollinearity was assessed using variance inflation factors. This process yielded age, sex, HER2 status, and signet-ring cell carcinoma for the clinicopathological model, and esophagogastric junction involvement,

Borrmann type IV, arterial-phase tumour attenuation, distant metastasis, and peritumoral clustered lymph nodes for the CT features model. The Integrated model included all nine predictors. All selected predictors were entered simultaneously, and no variables were removed on the basis of their multivariable P values.

Method S2. Model comparisons and internal validation

Internal validation used 1,000 nonparametric bootstrap resamples of 222 patients. In each resample, all coefficients were re-estimated using the fixed data-informed predictor set. Performance was evaluated in the bootstrap sample and in the original cohort, and mean optimism was subtracted from the apparent performance to obtain optimism-corrected AUCs, Brier scores, calibration intercepts, calibration slopes, and calibration curves. Differences between the optimism-corrected AUC estimates were used to summarize incremental discrimination. Bootstrap coefficient distributions were summarized using the median, 2.5th and 97.5th percentiles, and the proportion of estimates retaining the direction of the original coefficient. For the Integrated model, the optimism-corrected calibration slope of 0.815 was used as a uniform shrinkage factor. All nine slope coefficients were multiplied by 0.815, and the intercept was re-estimated with the shrunken linear predictor entered as an offset to preserve calibration-in-the-large. The resulting coefficients and re-estimated intercept are reported in Supplementary Table S11. The preliminary predictor-screening process was not repeated within the bootstrap resamples.

Method S3. Expression-gradient, cross-definition, and subgroup analyses

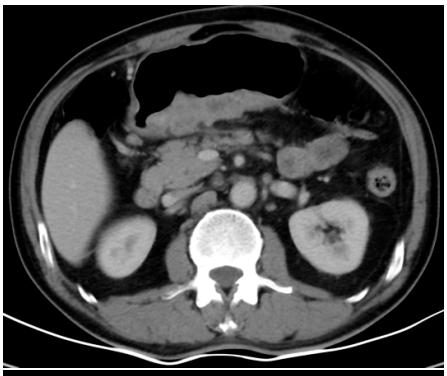
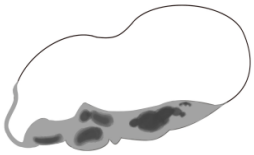
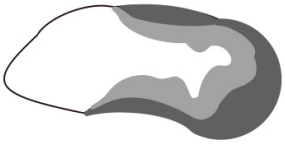
The Integrated model was developed using the primary ≥40% outcome. Following uniform shrinkage, its predicted probabilities were compared across the ordered <40%, 40–74%, and ≥75% expression strata using the Kruskal–Wallis test, with epsilon squared as the overall effect size. Pairwise differences were assessed using Dunn tests with Holm adjustment. For the comparison between the 40–74% and ≥75% groups, Cliff's delta was calculated with a 95% confidence interval based on 10,000 patient-level bootstrap resamples. Kendall's tau-b was used to assess ordinal associations of the shrunken model probability with the expression strata and recorded staining percentage. Patients were also divided into quintiles of the shrunken Integrated-model score, and trends in the observed prevalence of CLDN18.2 positivity were evaluated using the Cochran–Armitage trend test. The same shrunken score was evaluated against the ≥75% outcome without outcome-specific coefficient estimation. The full-cohort clinicopathological and CT features model scores and the shrunken Integrated-model score were also applied to the HER2-negative and HER2-positive subgroups without subgroup-specific coefficient estimation. The subgroup analyses were exploratory

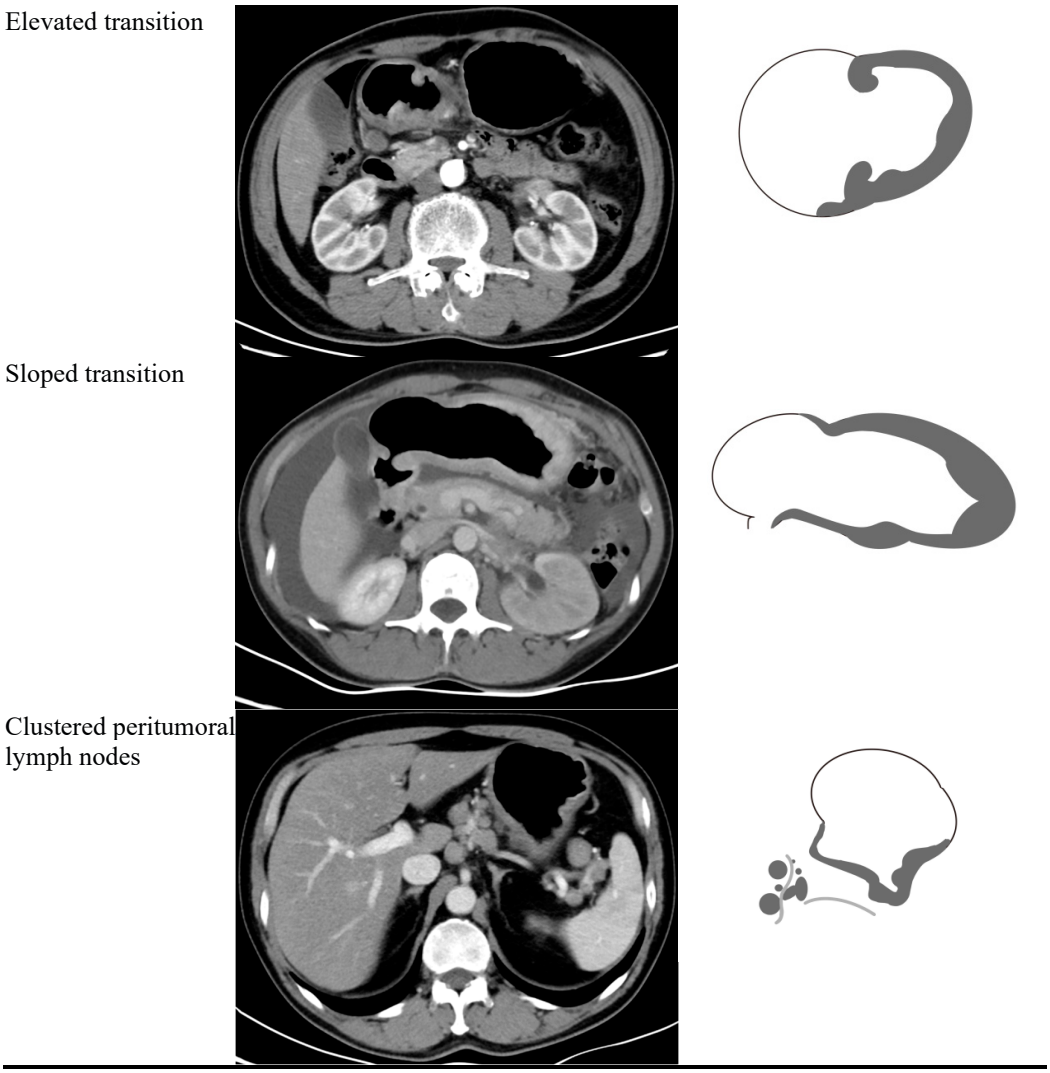
Method S4. Ordinal and continuous expression sensitivity analyses

Using the same nine predictors, a proportional-odds model analysed low expression (intensity <2 or percentage <40%), intermediate expression (intensity ≥2 and percentage 40–74%), and high expression (intensity ≥2 and percentage ≥75%). The proportional-odds assumption was evaluated using a cluster-robust global Wald comparison of the ≥40% and ≥75% cumulative-logit slopes. Recorded staining percentage divided by 100 was analysed using fractional logistic regression. Repeated stratified 10-fold cross-validation with 50

repeats was used; all coefficients were re-estimated in each training fold. Ordinal performance included macro one-versus-rest AUC, ordinal C-index, ranked probability score, expected-category mean absolute error, and quadratic-weighted kappa. Continuous performance included mean absolute error, root mean squared error, R^2, and Spearman correlation. The 2.5th and 97.5th percentiles describe variation across repeated cross-validation runs and are not confidence intervals. These sensitivity models were fitted independently and were not derived from the binary Integrated model or its uniform shrinkage adjustment.

Table S1. Representative CT images and measurement schematics of subjective imaging signs

Subjective imaging sign	CT image	Schematic diagram
Enhancement pattern Non-layered heterogeneous enhancement		
Layered heterogeneous enhancement		
Transition morphology		



Illustrative CT images and schematic diagrams of the subjective imaging signs evaluated in this study. The enhancement pattern was classified as non-layered or layered heterogeneous. Transition morphology was classified as elevated (defined by at least two abrupt angles) or sloped (showing a smooth interface with the adjacent gastric wall). Clustered lymph nodes were defined as the aggregation of at least five lymph nodes surrounding the primary tumour.

2. Supplementary Results

Equations of the final models for the ≥40% definition

The probability of CLDN18.2-positive status was calculated as $p(Y=1 \mid X)=1/[1+\exp(-\eta)]$, where η denotes the model-specific linear predictor. The three models were fitted in the full cohort using the primary ≥40% definition.

Clinicopathological model

$$\eta\text{clinicopathological} = 0.8059 - 1.0319\times(\text{age} > 60\ \text{years}) - 0.3770\times(\text{male sex}) - 0.9019\times(\text{HER2 positivity}) + 0.9185\times(\text{signet-ring cell carcinoma}).$$

CT features model

$$\eta\text{CT features} = -0.5702 - 0.8370\times(\text{EGJ involvement}) + 1.1604\times(\text{Borrmann type IV}) + 0.0118\times(\text{arterial-phase tumour attenuation, HU}) + 0.6376\times(\text{distant metastasis present}) - 0.7185\times(\text{peritumoural clustered lymph nodes}).$$

Integrated model

$$\eta\text{Integrated} = -0.8792 - 0.9103\times(\text{age} > 60\ \text{years}) + 0.2719\times(\text{male sex}) - 0.9691\times(\text{HER2 positivity}) + 0.8979\times(\text{signet-ring cell carcinoma}) - 0.7371\times(\text{EGJ involvement}) + 1.0931\times(\text{Borrmann type IV}) + 0.0142\times(\text{arterial-phase tumour attenuation, HU}) + 0.6863\times(\text{distant metastasis present}) - 0.3147\times(\text{peritumoural clustered lymph nodes}).$$

Uniformly shrunken Integrated model

$$\eta = -0.6681 - 0.7419\text{Xage} + 0.2216\text{Xmale} - 0.7898\text{XHER2} + 0.7318\text{Xsignet-ring} - 0.6007\text{XEGJ} + 0.8909\text{XBorrmann IV} + 0.0116\text{AHU} + 0.5593\text{Xmetastasis} - 0.2565\text{Xclustered LN}$$

$$\text{p(CLDN18.2} \geq 40\%) = 1 / [1 + \exp(-\eta)]$$

The uniformly shrunken Integrated-model probability was calculated as $p = 1/[1 + \exp(-\eta)]$, where η was derived from the re-estimated intercept and shrinkage-adjusted coefficients provided in Supplementary Table S11.

Variable coding and application

Binary predictors were coded as 1 for age >60 years, male sex, HER2-positive status, signet-ring cell carcinoma, esophagogastric junction involvement, Borrmann type IV

morphology, distant metastasis, and peritumoural clustered lymph nodes, and as 0 for their respective reference categories: age ≤60 years, female sex, HER2-negative status, absence of signet-ring cell carcinoma, no esophagogastric junction involvement, Borrmann type II or III morphology, no distant metastasis, and no peritumoural clustered lymph nodes. Arterial-phase tumour attenuation was entered continuously in Hounsfield units. For analyses requiring absolute Integrated-model probabilities, all nine original slope coefficients were multiplied by the optimism-corrected calibration slope of 0.815, and the intercept was re-estimated as −0.6681 to preserve calibration-in-the-large. The resulting uniformly shrunken model was applied to the expression-gradient and ≥75% analyses without outcome-specific coefficient re-estimation. EGJ, esophagogastric junction; HER2, human epidermal growth factor receptor 2; HU, Hounsfield unit; LN, lymph node.

Table S2. Interobserver agreement and Bland–Altman results for continuous CT measurements

Variable	Reader 1, mean ± SD	Reader 2, mean ± SD	ICC (95% CI)	Bias (R1–R2)	95% limits of agreement	Interpretation
Tumour long diameter (mm)	92.3 ± 46.5	91.5 ± 45.2	0.989 (0.985–0.991)	0.78	-12.63 to 14.18	Excellent
Tumour axial thickness (mm)	17.1 ± 6.2	17.1 ± 6.2	0.999 (0.998–0.999)	-0.02	-0.64 to 0.59	Excellent
Aortic CT attenuation, arterial phase (HU)	218.7 ± 55.1	219.0 ± 55.2	0.997 (0.997–0.998)	-0.27	-8.03 to 7.50	Excellent
Tumour CT attenuation, arterial phase (HU)	72.2 ± 23.8	73.1 ± 23.6	0.979 (0.972–0.984)	-0.91	-10.40 to 8.58	Excellent
NTER, arterial phase	0.35 ± 0.14	0.35 ± 0.14	0.985 (0.981–0.989)	-0.004	-0.050 to 0.041	Excellent
Aortic CT attenuation, venous phase (HU)	142.4 ± 20.1	142.8 ± 21.2	0.967 (0.957–0.975)	-0.43	-10.82 to 9.96	Excellent
Tumour CT attenuation, venous phase (HU)	79.2 ± 26.2	81.2 ± 25.7	0.953 (0.936–0.965)	-2.03	-17.14 to 13.09	Excellent
NTER, venous phase	0.56 ± 0.17	0.57 ± 0.16	0.940 (0.920–0.955)	-0.013	-0.122 to 0.096	Excellent

Variable	Reader 1, mean ± SD	Reader 2, mean ± SD	ICC (95% CI)	Bias (R1–R2)	95% limits of agreement	Interpretation
Venous-minus-arterial tumour attenuation difference (HU)	7.0 ± 17.3	8.1 ± 16.9	0.891 (0.860–0.916)	-1.12	-16.66 to 14.43	Good
Largest lymph-node short-axis diameter (mm)	11.2 ± 9.9	11.2 ± 9.9	1.000 (0.999–1.000)	-0.00	-0.55 to 0.54	Excellent

ICC values were estimated using two-way random effects, absolute agreement, and a single-measure model [ICC(2,1)]. The Bland-Altman bias was calculated as reader 1 minus Reader 2. CI, confidence interval; CT, computed tomography; HU, Hounsfield unit; ICC, intraclass correlation coefficient; LN, lymph node; NTER, normalised tumour enhancement ratio. Non-contrast tumour and aortic attenuation values are not included because paired Reader 2 measurements were unavailable.

Table S3. Interobserver agreement for categorical CT features

Variable	Type	Agreement, %	Kappa (95% CI)[a]	Interpretation[b]
Tumor location	Nominal (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Involved gastric region	Ordinal (quadratically weighted kappa)	99.1	0.990 (0.977–1.000)	Almost perfect
Circumferential involvement	Ordinal (quadratically weighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
T stage	Ordinal (quadratically weighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Borrmann type	Nominal (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Enhancement homogeneity	Nominal (unweighted kappa)	95.9	0.938 (0.898–0.978)	Almost perfect
Enhancement degree	Ordinal (quadratically weighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Enhancement pattern	Nominal (unweighted kappa)	97.7	0.963 (0.930–0.995)	Almost perfect
N stage	Ordinal (quadratically weighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
EGJ involvement	Binary	100.0	1.000 (1.000–	Almost perfect

Variable	Type	Agreement, %	Kappa (95% CI)[a]	Interpretation[b]
	(unweighted kappa)		1.000)	
Transition morphology (sloped versus elevated)	Binary (unweighted kappa)	98.6	0.916 (0.821–1.000)	Almost perfect
Peritumoral clustered LNs	Binary (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Peak enhancement phase	Binary (unweighted kappa)	92.3	0.788 (0.692–0.884)	Substantial
Large LN (>3 cm)	Binary (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Heterogeneous LN enhancement	Binary (unweighted kappa)	97.7	0.937 (0.882–0.991)	Almost perfect
Confluent LNs	Binary (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Retroperitoneal LNs	Binary (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Peritoneal metastasis	Binary (unweighted kappa)	96.8	0.932 (0.882–0.982)	Almost perfect
Ascites	Binary (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect
Distant metastasis	Binary (unweighted kappa)	100.0	1.000 (1.000–1.000)	Almost perfect

Quadratically weighted Cohen kappa was used for ordinal variables; unweighted Cohen kappa was used for binary and nominal multicategory variables.

Kappa values were interpreted as poor (<0), slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), or almost perfect (0.81–1.00).

CI, confidence interval; CT, computed tomography; EGJ, esophagogastric junction; LN, lymph node.

Table S4. Bootstrap-corrected performance of the three models

Model	Apparent AUC	Corrected AUC	Apparent Brier	Corrected Brier	Corrected intercept	Corrected slope	Uniform shrinkage factor	Re-estimated intercept

Model	Apparent AUC	Corrected AUC	Apparent Brier	Corrected Brier	Corrected intercept	Corrected slope	Uniform shrinkage factor	Re-estimated intercept
Clinicopathological	0.724	0.706	0.208	0.218	0.025	0.912	-	-
CT features	0.722	0.698	0.209	0.221	0.035	0.870	-	-
Integrated	0.773	0.735	0.189	0.208	0.051	0.815	0.815	-0.668

Table S5. Incremental model comparisons

Comparison	Apparent ΔAUC	Paired-bootstrap 95% CI for apparent ΔAUC	Difference between optimism-corrected AUC estimates
Integrated vs clinicopathological	0.049	0.006–0.091	0.029
Integrated vs CT features	0.051	0.006–0.100	0.037
CT features vs clinicopathological	−0.002	−0.084–0.073	−0.008

In every bootstrap sample, the complete final candidate set was refitted without variable selection. All 1,000 resamples were successfully analysed. Detailed optimism components are reported for transparency. AUC, area under the receiver operating characteristic curve; CI, confidence interval.

Table S6. Integrated-model probabilities across CLDN18.2 expression strata

Panel A. Descriptive statistics

Expression stratum	n	Shrunken Integrated-model score, median (IQR)	Shrunken predicted probability, median (IQR)
<40%	92	−0.111 (−0.738 to 0.423)	0.472 (0.323–0.604)
40–74%	30	0.318 (−0.349 to 0.976)	0.579 (0.414–0.726)
≥75%	100	0.930 (0.204–1.583)	0.717 (0.551–0.830)

Panel B. Overall differences and ordered associations

Analysis	Statistic/effect	P value
Kruskal–Wallis difference across three strata	$H=56.711$; $\varepsilon^2=0.250$	4.85×10^{-13}
Shrunken Integrated-model score vs ordered expression stratum	Kendall $\tau b=0.401$	2.80×10^{-14}
Shrunken Integrated-model score vs recorded staining percentage	Kendall $\tau b=0.389$	1.03×10^{-15}
Effect size, 40–74% vs ≥75%	Cliff's $\delta = -0.386$ (95% bootstrap CI, −0.589 to −0.163)	-

Panel C. Pairwise Dunn tests

Comparison	Z	Unadjusted P	Holm-adjusted P
<40% vs 40–74%	−2.222	0.0263	0.0263
<40% vs ≥75%	−7.514	5.73×10^{-14}	1.72×10^{-13}
40–74% vs ≥75%	−2.970	0.0030	0.0060

Probabilities were calculated using the uniformly shrunken Integrated model. All nine slope coefficients were multiplied by 0.815, and the intercept was re-estimated to −0.6681 to preserve calibration-in-the-large. No outcome-specific coefficient estimation was performed for the expression-gradient or ≥75% analyses.

Table S7. Observed CLDN18.2 positivity across quintiles of the shrunken Integrated-model score

Quintile	n	Shrunken Score range	≥40% positive, n (%) [95% CI]	≥75% positive, n (%) [95% CI]
Q1	45	−2.766 to −0.422	10 (22.2) [12.5–36.3]	3 (6.7) [2.3–17.9]
Q2	44	−0.401 to 0.106	18 (40.9) [27.7–55.6]	12 (27.3) [16.3–41.8]
Q3	44	0.106 to 0.600	30 (68.2) [53.4–80.0]	24 (54.5) [40.1–68.3]
Q4	44	0.609 to 1.328	34 (77.3) [63.0–87.2]	27 (61.4) [46.6–74.3]
Q5	45	1.339 to 2.947	38 (84.4) [71.2–92.3]	34 (75.6) [61.3–85.8]

The Cochran–Armitage trend statistics were Z=6.905 ($P=5.01\times10^{-12}$) for the ≥40% definition and Z=7.312 ($P=2.64\times10^{-13}$) for the ≥75% definition. Confidence intervals are Wilson intervals.

Exploratory HER2 subgroup analyses

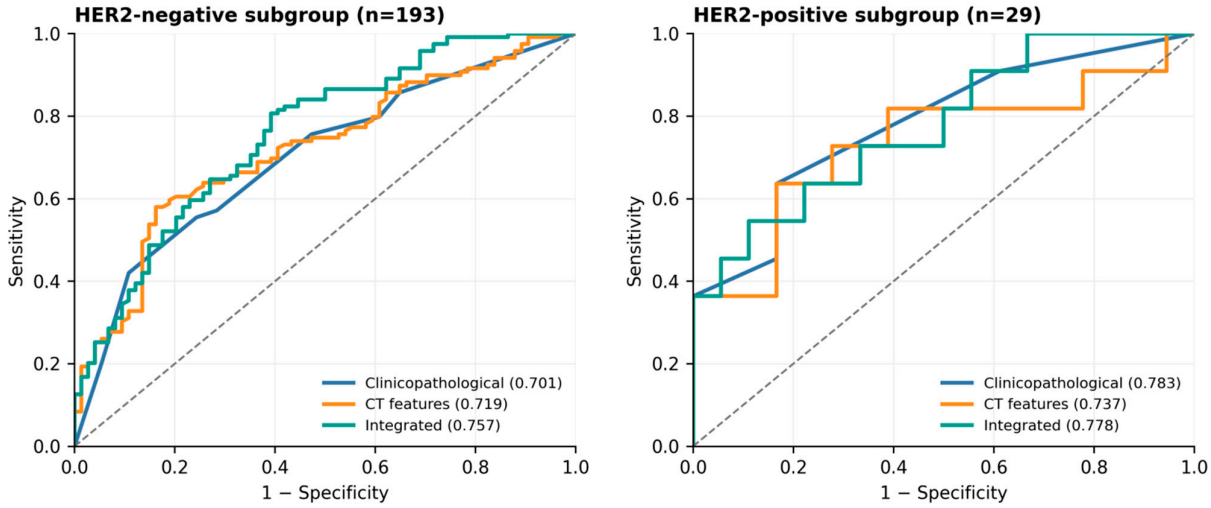
In the HER2-negative subgroup, the clinicopathological, CT features, and Integrated models yielded AUCs of 0.701, 0.719, and 0.757, respectively, without subgroup-specific refitting. In the HER2-positive subgroup, the corresponding AUCs were 0.783, 0.737, and 0.778. The unchanged Integrated model yielded an AUC of 0.804 for the ≥75% outcome in HER2-positive patients. However, this subgroup comprised only 29 patients, including 11 ≥40% events and 8 ≥75% events, resulting in wide confidence intervals (Table S8 and Figure S1).

Table S8. Cross-definition discrimination and exploratory HER2 subgroup analyses

Population	Outcome/model	n	Events	AUC (95% bootstrap CI)

Population	Outcome/model	n	Events	AUC (95% bootstrap CI)
Full cohort	Integrated model, ≥75% outcome	222	100	0.781 (0.718–0.836)
HER2-negative	Clinicopathological, ≥40% outcome	193	119	0.701 (0.633–0.775)
HER2-negative	CT features, ≥40% outcome	193	119	0.719 (0.645–0.791)
HER2-negative	Integrated, ≥40% outcome	193	119	0.757 (0.684–0.830)
HER2-negative	Integrated model, ≥75% outcome	193	92	0.767 (0.699–0.829)
HER2-positive	Clinicopathological, ≥40% outcome	29	11	0.783 (0.586–0.943)
HER2-positive	CT features, ≥40% outcome	29	11	0.737 (0.511–0.927)
HER2-positive	Integrated, ≥40% outcome	29	11	0.778 (0.561–0.929)
HER2-positive	Integrated, ≥75% outcome	29	8	0.804 (0.589–0.971)

The full-cohort model scores were applied without subgroup-specific coefficient estimation. The uniformly shrunken Integrated-model score was used for analyses involving absolute Integrated-model probabilities.



Supplementary Figure S1. Discrimination in HER2-defined subgroups.

Receiver operating characteristic curves for the clinicopathological, CT features, and Integrated models in the HER2-negative and HER2-positive subgroups. Full-cohort coefficients were applied without subgroup-specific refitting. Estimates in the HER2-positive subgroup are exploratory because only 29 patients were available. AUC, area under the receiver operating characteristic curve; HER2, human epidermal growth factor receptor 2.

Ordinal and continuous expression sensitivity analyses

The associations between the nine predictors and higher CLDN18.2 expression did not differ materially between the ⩾40% and ⩾75% thresholds, supporting the proportional-odds assumption (global Wald χ^2=5.08, 9 df, P=0.827). In repeated cross-validation, the ordinal model showed modest overall discrimination across the ordered expression categories, with a mean macro one-versus-rest AUC of 0.669, an ordinal C-index of 0.718, and a ranked probability score of 0.206. The class-specific one-versus-rest AUCs were 0.726 for low expression, 0.535 for intermediate expression, and 0.746 for high expression, indicating that discrimination was concentrated in the low- and high-expression categories, whereas the intermediate category was poorly distinguished from the other two categories. For the continuous staining percentage, the predicted and recorded values showed a moderate rank correlation (Spearman correlation, 0.466), but the model explained only 21.7% of the variation in the recorded percentages (R^2=0.217). Similar results were obtained when the ordinal categories were defined using staining percentage alone. Overall, the nine-predictor profile provided modest information on the ordering of CLDN18.2 expression, but it did not reliably identify the intermediate-expression category or estimate the exact staining percentage of an individual tumour. (Tables S9–S10 and Figure S2).

Table S9. Ordinal and fractional logistic expression models

Model	Predictor	Estimate (95% CI)	P value
Ordinal	Age >60 years	0.47 (0.25–0.86)	0.014
Ordinal	Male sex	1.25 (0.64–2.46)	0.513
Ordinal	HER2 positive	0.40 (0.16–0.95)	0.037
Ordinal	Signet-ring cell carcinoma	1.98 (1.08–3.62)	0.028
Ordinal	EGJ involvement	0.45 (0.24–0.86)	0.015
Ordinal	Borrmann type IV	3.45 (1.55–7.67)	0.002
Ordinal	Arterial-phase CT value (per 10 HU)	1.18 (1.04–1.35)	0.012

Model	Predictor	Estimate (95% CI)	P value
Ordinal	Distant metastasis	2.28 (1.29–4.02)	0.004
Ordinal	Peritumoral clustered LNs	0.64 (0.32–1.28)	0.206
Fractional logistic	Age >60 years	0.58 (0.37–0.93)	0.024
Fractional logistic	Male sex	1.24 (0.75–2.04)	0.400
Fractional logistic	HER2 positive	0.51 (0.26–0.99)	0.047
Fractional logistic	Signet-ring cell carcinoma	2.01 (1.30–3.13)	0.002
Fractional logistic	EGJ involvement	0.46 (0.28–0.75)	0.002
Fractional logistic	Borrmann type IV	2.39 (1.40–4.10)	0.002
Fractional logistic	Arterial-phase CT value (per 10 HU)	1.14 (1.03–1.26)	0.009
Fractional logistic	Distant metastasis	1.69 (1.09–2.60)	0.018
Fractional logistic	Peritumoral clustered LNs	0.64 (0.36–1.15)	0.135

For the ordinal model, estimates are common odds ratios for membership in a higher expression category. For the fractional logistic model, estimates are odds ratios for the conditional mean expression proportion. All nine predictors were entered simultaneously. The proportional-odds global Wald test was χ^2=5.08 with 9 df (P=0.827). CI, confidence interval.

Table S10. Performance from 50 repeated stratified 10-fold cross-validation runs

Model	Metric	Mean	2.5th–97.5th percentiles
Ordinal proportional-odds	Macro one-vs-rest AUC	0.669	0.656–0.682
Ordinal proportional-odds	Low expression vs rest AUC	0.726	0.713–0.740
Ordinal proportional-odds	Intermediate expression vs rest AUC	0.535	0.508-0.562
Ordinal proportional-odds	High expression vs rest AUC	0.746	0.734-0.762
Ordinal proportional-odds	Ordinal C-index	0.718	0.709–0.732
Ordinal proportional-odds	Ranked probability score	0.206	0.200–0.211
Ordinal proportional-odds	MAE of expected category (0-2)	0.719	0.707–0.729
Fractional logistic	MAE of expression proportion	0.286	0.281–0.291
Fractional logistic	RMSE of expression proportion	0.341	0.335–0.348
Fractional logistic	R-squared	0.217	0.188–0.246
Fractional logistic	Spearman correlation	0.466	0.440–0.491
Binary logistic ≥75%	AUC	0.741	0.726–0.756

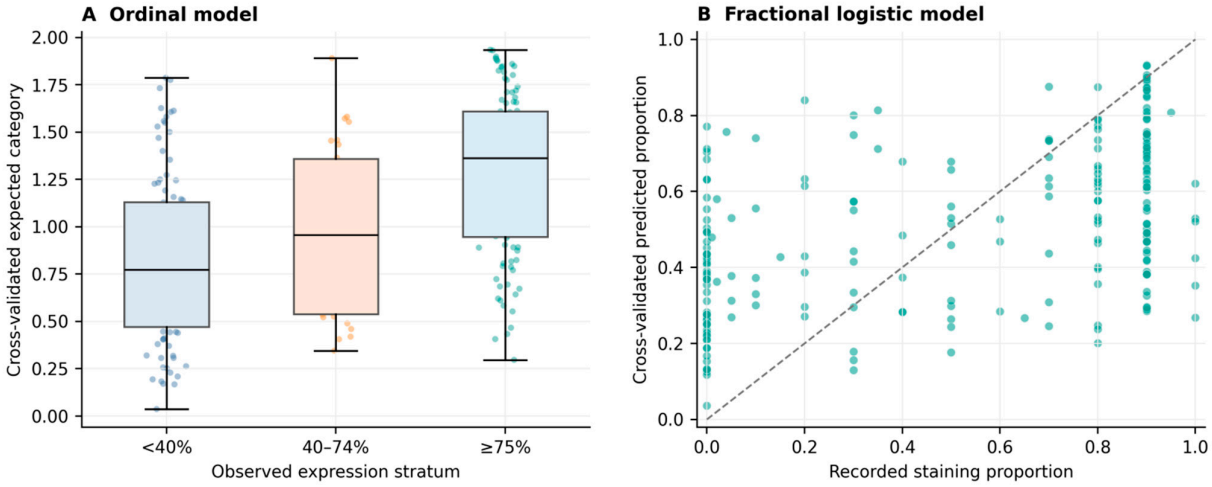
Model	Metric	Mean	2.5th–97.5th percentiles
Binary logistic ≥75%	Brier score	0.207	0.200–0.213
Binary logistic ≥75%	Calibration slope	0.752	0.688–0.827
Percentage-only ordinal sensitivity	Macro one-vs-rest AUC	0.670	0.659–0.683
Percentage-only ordinal sensitivity	Ordinal C-index	0.721	0.711–0.735

Every model coefficient was re-estimated within each training fold. The percentiles describe variation across the 50 repeated cross-validation runs and are not confidence intervals. AUC, area under the receiver operating characteristic curve; MAE, mean absolute error; RMSE, root mean squared error.

Table S11. Uniformly shrunken coefficients and re-estimated intercepts for the prediction models

Model term	Original coefficient	Shrinkage factor	Final coefficient
Intercept	-0.8792	Not applied	-0.6681
Age >60 years	-0.9103	0.815	-0.7419
Male sex	0.2719	0.815	0.2216
HER2 positivity	-0.9691	0.815	-0.7898
Signet-ring cell carcinoma	0.8979	0.815	0.7318
Esophagogastric junction involvement	-0.7371	0.815	-0.6007
Borrmann type IV	1.0931	0.815	0.8909
Arterial-phase tumour attenuation, per 1 HU	0.0142	0.815	0.0116
Distant metastasis present	0.6863	0.815	0.5593
Peritumoural clustered lymph nodes	-0.3147	0.815	-0.2565

The optimism-corrected calibration slope of 0.815 was applied uniformly to all nine slope coefficients. The original intercept was not multiplied by the shrinkage factor; it was re-estimated to preserve calibration-in-the-large. The final model probability is calculated as $p = 1/[1 + \exp(-\eta)]$, using the final coefficients shown above. Coefficients are reported on the log-odds scale.



Supplementary Figure S2. Cross-validated ordinal and continuous expression predictions.

(A) Expected expression category from the proportional-odds model across observed expression strata. (B) Fractional-logistic predicted expression proportion versus the recorded staining proportion. Predictions are patient-level out-of-fold values averaged across 50 repeated stratified 10-fold cross-validation runs.