npj | precision oncology

Article

Published in partnership with The Hormel Institute, University of Minnesota

https://doi.org/10.1038/s41698-025-01224-w

Mucin phenotype-based deep learning framework for intestinal metaplasia-carcinogenesis progression prediction

Xiaoyang Wu[1,2], Fang Wang[1,3], Weiyou Dai[4], Chuxuan Ni[1,2], Liping Sun[1,2], Yuehua Gong[1,2], Nannan Dong[1,2], Zeyang Wang[1,2], Liang Li[1,2], Qian Xu[1,2], Jingjing Jing[1,2], Shixuan Shen[1,2], Huakang Tu[4] ✉ & Yuan Yuan[1,2] ✉

This study decodes spatiotemporal mucin dynamics in gastric carcinogenesis, revealing gastric-type markers (MUC5AC/MUC6) decline progressively while intestinal-type markers (MUC2/CD10) peak in gastric intestinal metaplasia (GIM) before decreasing in gastric cancer (GC). We developed MPMR, a dual-function UNI-pretrained Vision Transformer (ViT) model, which directly predicts four mucin markers from H&E whole-slide images with near-perfect accuracy (AUC: 0.921–0.997) and generates interpretable simulated staining heatmaps via adversarial learning. Integrating these outputs with clinical variables, the MPMR-IMCP risk model significantly outperformed clinical-only models (ΔAUC = 0.050), enabling both phenotype analysis and GIM risk stratification without specialized staining. Validated in longitudinal cohorts from Chinese GC high-incidence regions, this framework offers an efficient solution for monitoring GIM malignant transformation.

Gastric cancer (GC) is the fifth most common malignancy worldwide and the fifth leading cause of cancer-related deaths[1]. The epidemiological distribution of GC exhibits substantial geographical variations, with China representing a high-incidence region where northern areas demonstrate an incidence rate 2.6-fold higher than the national average[2]. The pathogenesis of GC follows the well-established Correa cascade progression model: superficial gastritis (GS) → atrophic gastritis (GA) with gastric intestinal metaplasia (GIM) → gastric dysplasia (GD) → GC[3,4]. Lauren classification categorizes GC predominantly into intestinal-type and diffuse-type subtypes, with intestinal-type accounting for the majority of cases[5]. Within this pathological continuum, GIM represents the most clinically significant precancerous stage, in which the risk of malignant transformation is three times greater for the incomplete subtype than for the complete type[6]. Therefore, accurate GIM risk stratification is critical important for targeted interventions. However, unfortunately, current stratification standards, such as the operative link for gastric intestinal metaplasia assessment (OLGIM)[7,8] are still limited to quantitative assessment, while ignoring the role of key molecular typing, such as mucin phenotype.

Mucin, as the fundamental molecular component for maintaining the function of the mucosal barrier, exhibit diverse physiological roles[9]. Gastric mucins (MUC5AC and MUC6) protect the epithelial surface from acid and enzymatic digestion through formation of a protective mucus-gel layer[10], while intestinal mucins (MUC2) and CD10 are involved in mucosal remodeling within the metaplastic region[11,12]. Together, MUC5AC/MUC6 (gastric phenotype markers) and MUC2/CD10 (intestinal phenotype markers) constitute a complex mucin phenotype system in gastric mucosa, demonstrating distinct expression patterns across different GIM subtypes that enable discrimination between complete and incomplete GIM. This molecular classification carries considerable clinical promise in the risk stratification of GIM, given that the malignant potential varies considerably among the different subtypes. While image-enhanced endoscopy can detect the presence of GIM[13], it primarily relies on morphological patterns and lacks the capability to differentiate between complete and incomplete GIM subtypes. Although numerous studies have established the link between mucin expression patterns and gastric disease progression[14–16], the majority have been constrained by several limitations, such as reliance on cross-sectional analyses at single disease stages and absence of longitudinal dynamic data spanning from precancerous lesions to malignant transformation, thus failing to characterize the phenotypic transformation of the entire Correa cascade. Furthermore, traditional immunohistochemical (IHC) staining methods still rely on manual interpretation, which has three main drawbacks, including low analytical throughput, significant

[1]Tumor Etiology and Screening Department of Cancer Institute, and Key Laboratory of Cancer Etiology and Prevention in Liaoning Education Department, the First Hospital of China Medical University, Shenyang, Liaoning, China. [2]Key Laboratory of Gastrointestinal Cancer Etiology and Prevention, Shenyang, Liaoning, China. [3]Department of Pathology, Qilu Hospital of Shandong University Dezhou Hospital, Dezhou, Shandong, China. [4]Center of Clinical Big Data and Analytics of the Second Affiliated Hospital and School of Public Health, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. ✉e-mail: huakangtu@zju.edu.cn; yuanyuan@cmu.edu.cn

translation subjectivity, and huge cost burden. These jointly limit the clinical application potential of mucin phenotype prediction in the stratification assessment of GIM carcinogenesis risk.

The integration of digital pathology with deep learning technologies has revolutionized this research predicament. Several studies have demonstrated the technical feasibility of predicting IHC staining patterns directly from routine hematoxylin and eosin (H&E)-stained images using deep learning algorithms[17,18]. This computational approach provides at least three key advantages including significant reduction in pathological diagnostic costs, streamlined clinical workflows and preservation of valuable tissue specimens, while simultaneously enabling objective and quantitative analysis[18–22]. Vision Transformers (ViT), representing a major advancement in deep learning architectures, employ self-attention mechanisms to effectively model long-range spatial relationships within whole-slide images (WSIs), thereby overcoming the inherent limitations of convolutional neural networks (CNNs) in global feature representation[23–25]. In medical imaging applications, ViT-based frameworks including UNI, CONCH, and TITAN have made substantial progress in critical pathological tasks such as tissue classification and tumor detection[26–28]. However, despite these technological advancements, they are currently mostly used in the research of tumor prognosis and therapeutic effect prediction[29,30], and have little involvement in the prediction of carcinogenesis risk of precancerous diseases. The potential application of ViT models based on mucin typing for risk stratification and malignant progression prediction in gastric precancerous lesions - particularly GIM - has not yet been investigated.

In this study, a long-term follow-up cohort and biological specimen bank established in the high-incidence areas of GC in northern China were adopted[31,32]. We systematically investigated the expression patterns and classifications of mucin phenotypic biomarkers (MUC5AC, MUC6, MUC2, and CD10) throughout the dynamic progression of gastric diseases. Further, by combining full-section imaging with advanced deep learning techniques, we have developed two innovative models, including the Mucin Phenotypic Marker Recognizer (MPMR) and its clinical application extension, the MPMR-Intestinal Metaplasia Carcinogenesis Predictor (MPMR-IMCP). The MPMR model, built upon ViT architecture, incorporates pre-trained weights from the UNI pathology foundation model to achieve precise identification of mucin expression patterns directly from H&E-stained WSIs, while also extracting quantitative phenotypic features and introduced simulated staining heat maps as diagnostic visualization tools. The MPMR-IMCP model extends this capability by integrating these deep learning-derived mucin biomarkers to achieve dual functions including GIM subtypes classification and their risk prediction of malignant transformation. Overall, this comprehensive system presents a new framework for precisely assessing the risk of GIM malignant transformation based on molecular typing rather than the previous simple quantification alone.

Results

Study cohort baseline characteristics

Four independent cohorts were included in this study, with baseline characteristics summarized in Table 1. For the analysis of IHC staining, a cohort of 288 patients with gastric diseases from the FHCMU-IHC resource platform at the First Hospital of China Medical University was utilized. For the MPMR model development phase, two histopathological image datasets were employed: (1) an internal dataset (FHCMU-ESD) from The First Hospital of China Medical University, consisting of 74 cases used for model training and internal validation; and (2) an external dataset (ZH-EGD) from a high-risk Gastric Diseases Screening Program in Zhuanghe (ZH-GDSP), Liaoning Province, comprising 78 cases used for independent external validation to evaluate model generalizability. Each case included one H&E-stained WSIs and four specially stained WSIs targeting distinct mucin phenotypic markers, resulting in a total of 760 pathological images. For the MPMR-IMCP model development, the ZH-GDSP contributed 240 samples, including 80 cases with subsequent carcinogenesis and 160 age- and sex-matched controls (case-to-control ratio 1:2). The cases in the ZH-GDSP cohort were categorized into a non-cancer group and a cancer group. All cases included one H&E-stained WSIs and integrated multidimensional data, including demographic characteristics, pathological diagnoses, serological markers, and clinical features. All cases included a confirmed underlying pathological diagnosis of GIM.

Expression characteristics and trends of mucin phenotypic biomarkers in various gastric diseases

IHC staining simultaneously detected the expression of four mucin phenotype markers: MUC5AC, MUC6, MUC2, and CD10 proteins in various gastric disease tissues. MUC5AC staining was predominantly located in the superficial foveolar epithelium, while MUC6 was mainly distributed in the pyloric gland and fundic gland neck mucous cells. MUC2 staining concentrated in goblet cells, and CD10 was primarily observed surrounding the brush border. The expression levels of these indicators are detailed in the upper part of Supplementary Table 2 and Fig. 1. In the stepwise progression proceeds through GS, GA, Low-grade Intraepithelial Neoplasia (LGIN), and High-grade Intraepithelial Neoplasia (HGIN), to intestinal-type GC, the expression level of MUC5AC gradually decreased ($P_{GS\text{-}GA} = 0.015$, $P_{GA\text{-}LG} < 0.001$, $P_{LG\text{-}HG} < 0.001$), remaining relatively stable at the high-grade intraepithelial neoplasia and GC stages (Fig. 2a–f). The expression level of MUC6 also gradually decreased ($P_{GS\text{-}GA} = 0.002$, $P_{GA\text{-}LG} < 0.001$), stabilizing after the low-grade intraepithelial neoplasia stage (Fig. 2g–l). MUC2 initially rose continuously, then decreased at the GC stage ($P_{GS\text{-}GA} < 0.001$, $P_{HG\text{-}GC} < 0.001$) (Fig. 2m–r). CD10 exhibited fluctuating changes in expression ($P_{GS\text{-}GA} < 0.001$) (Fig. 2s–x).

Table 1 | Patient baseline table

Clinicopathological features		FHCMU-IHC	FHCMU-ESD	ZH-EGD	ZH-GDSP	
					Non-cancerous	cancerous
Patients		288	74	78	160	80
H&E slides		0	74	78	160	80
IHC slides		1152	296	312	0	0
Age	<50	92	14	27	90	18
	≥50	196	60	51	70	62
Gender	Male	176	22	53	91	68
	Female	112	52	25	69	12

In the FHCMU-IHC cohort, IHC scoring involved different disease categories and was further subdivided according to disease type, including 26 cases of GS, 71 cases of GA, 39 cases of LGIN, 53 cases of HGIN, and 99 cases of GC. All cases in the cohorts presented with a background lesion of GIM. *GIM* Gastric Intestinal metaplasia, *GS* Superficial Gastritis, *GA* Atrophic Gastritis, *LGIN* Low-grade Intraepithelial Neoplasia, *HGIN* High-grade Intraepithelial Neoplasia, *GC* Gastric Cancer.

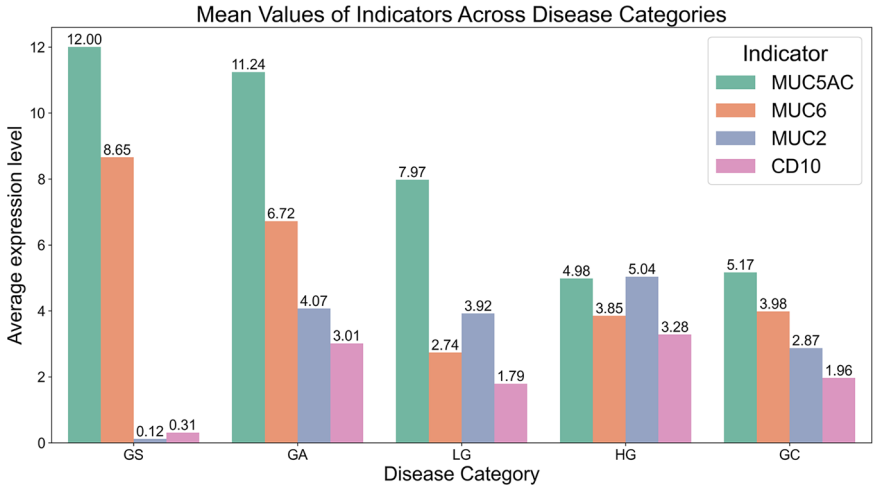

Fig. 1 | Expression levels of mucin phenotype biomarkers in different gastric diseases. The expression levels of different mucin phenotypic markers in various gastric diseases. GS Superficial Gastritis, GA Atrophic Gastritis, LG Low-grade Intraepithelial Neoplasia, HG High-grade Intraepithelial Neoplasia, GC Gastric Cancer.

Classification characteristics and trends of mucin phenotypes in different gastric diseases

Furthermore, we classified the mucin phenotypes based on the expression of mucin phenotypic markers into non-intestinal gastric mucin, complete intestinal type I, incomplete intestinal type II, and type III (see Supplementary Table 1). We examined the protein expression of different mucin phenotypes in gastric diseases, with specific expression levels detailed in the latter part of Supplementary Table 2 and in Fig. 3a. The results showed that during the progression of the disease, non-intestinal type gastric mucin gradually decreased, reaching its lowest level during high-grade intraepithelial neoplasia and GC stages ($P_{GS\text{-}GA} < 0.001$, $P_{GA\text{-}LG} < 0.001$, $P_{LG\text{-}HG} = 0.036$) (Fig. 3b); complete intestinal type I initially increased then decreased ($P_{GS\text{-}GA} < 0.001$, $P_{LG\text{-}HG} = 0.040$, $P_{HG\text{-}GC} < 0.001$) (Fig. 3c); both incomplete intestinal types II ($P_{GA\text{-}LG} < 0.001$) and III ($P_{GS\text{-}GA} < 0.001$, $P_{GA\text{-}LG} < 0.001$, $P_{HG\text{-}GC} = 0.015$) exhibited a pattern of initially increasing followed by decreasing, stabilizing after GA as shown in the trend graphs (Fig. 3d, e).

MPMR training and internal validation performance

The MPMR model dataset was constructed based on region-of-interest (ROI) annotations. Pathology experts comprehensively reviewed all available H&E-stained images and mucin marker-stained slides, with ROIs annotated using QuPath[33] software. Annotated regions were divided into 224 × 224-pixel tiles, yielding a total of 411,206 labeled tiles. All tiles underwent standardized preprocessing steps, including random rotation, flipping, color jittering, and normalization. Tile distribution details are presented in Table 2. The MPMR model, built on a ViT-based architecture with UNI-2 pretrained weights, analyzed these tiles to identify expression patterns of MUC5AC, MUC6, MUC2, and CD10 in gastric mucosal tissues. The complete workflow is illustrated in Fig. 4b, c.

In internal validation using the FHCMU-ESD cohort, the MPMR model demonstrated strong performance in recognizing mucin markers. The area under curve (AUC) values of receiver operating characteristic (ROC) curve were 0.983 (95% CI 0.982–0.985) for MUC5AC, 0.949 (95% CI 0.944–0.954) for MUC6, 0.984 (95% CI 0.981–0.984) for MUC2, and 0.998 (95% CI 0.996–1.000) for CD10 (Fig. 5). To assess the discriminative ability of mucin phenotypic markers, confusion matrices were generated, and trends in accuracy, sensitivity, and specificity throughout the training process were recorded (Supplementary Fig. 1a–h). The precision, recall, and F1 score for each label were also calculated, with detailed statistics for this section provided in the upper half of Table 3.

Independent External Validation Performance of MPMR

Following the completion of MPMR model training and internal validation, we conducted external validation using the ZH-EGD cohort, applying the same preprocessing procedures as those used in the FHCMU-ESD study. The best-performing MPMR model from the internal validation phase was applied directly without any further adjustments to evaluate its real-world performance. In the ZH-EGD validation set, the AUC values for mucin marker recognition were as follows: 0.994 (95% CI 0.994–0.996) for MUC5AC, 0.997 (95% CI 0.997–0.998) for MUC6, 0.980 (95% CI 0.978–0.984) for MUC2, and 0.921 (95% CI 0.911–0.931) for CD10 (Fig. 6). In addition to AUC metrics, sensitivity, specificity, precision, and F1 scores were calculated to provide a comprehensive assessment of external validation performance. Confusion matrices for the validation sets are shown in Supplementary Fig. 1i–l, with detailed statistical information summarized in the lower half of Table 3.

Feature visualization and simulated staining heatmaps of MPMR

To validate the effectiveness of image feature extraction and the accuracy of mucin phenotypic marker recognition, this study employed two visualization methods, including t-distributed stochastic neighbor embedding (t-SNE)[34] and simulated staining heatmaps, to assess the discriminative capability of the MPMR model. First, the 1536-dimensional features extracted by the model were reduced to 50 dimensions using principal component analysis (PCA), followed by t-SNE mapping to a two-dimensional space (Fig. 7a–d). The results showed clear separation between positive and negative tiles for all four markers in the 2D space, confirming the model's ability to capture characteristic differences between categories. Figure 7e–l displays comparative results between simulated staining heatmaps and actual IHC staining images: MPMR-predicted MUC5AC expression regions were primarily distributed in the superficial foveolar epithelium, MUC6 expression was concentrated in pyloric glands and mucous neck cells of the gastric antrum, MUC2 showed specific localization to goblet cells, and CD10 was mainly expressed in brush border membranes. These predictions exhibited high concordance with IHC staining characteristics, and the simulated heatmaps showed strong agreement with interpretations by pathological experts, confirming the model's capability to accurately predict mucin marker expression patterns and spatial distributions using H&E images alone. In addition, simulated staining heatmaps were generated for gastric endoscopic biopsy specimens to demonstrate the model's predictive capability across different types of specimens (Supplementary Fig. 2).

Construction of the MPMR-IMCP

Quantitative analysis of H&E-stained images using the MPMR model provided mucin phenotypic parameters. For each case, we calculated the absolute quantities and relative proportions of four mucin markers (MUC5AC, MUC6, MUC2, CD10), enabling GIM phenotypic classification based on mucin typing (Supplementary Table 1), 7 mucin phenotype characteristic parameters were extracted. In addition, we

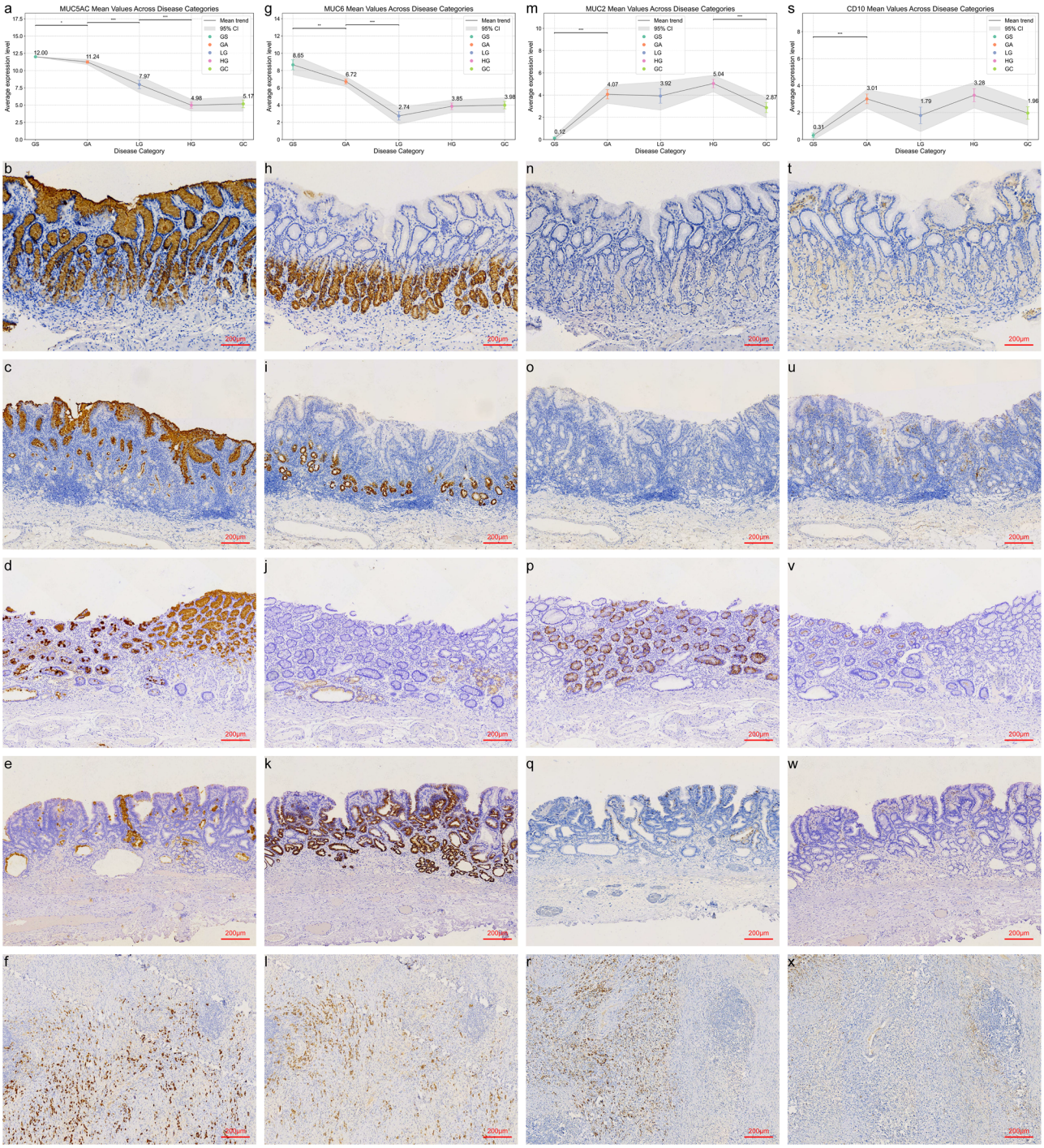

Fig. 2 | Trends of mucin phenotype biomarkers in different gastric diseases. The trends of different mucin phenotypic markers in various gastric diseases, as well as the IHC staining of each indicator. **a**–**f** Expression trend of MUC5AC and its IHC staining in different types of diseases. **g**–**l** MUC6. **m**–**r** MUC2. **s**–**x** CD10. GS Superficial Gastritis, GA Atrophic Gastritis, LG Low-grade Intraepithelial Neoplasia, HG High-grade Intraepithelial Neoplasia, GC Gastric Cancer. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

extracted 11 clinical parameters from the ZH-GDSP cohort (240 cases), including age, sex, pathological diagnosis, five serological indices (pepsinogen I [PGI], pepsinogen II [PGII], PGI/II ratio, Helicobacter pylori infection status [HP], and gastrin-17 [G17]), smoking history, alcohol consumption history, and family cancer history. These clinical features were combined with mucin phenotypic parameters to construct the MPMR-IMCP model.

Systematic comparisons of different feature combinations were performed (Table 4), with detailed statistical data presented in Supplementary Table 3 and a visualization of the feature selection process shown in Supplementary Figs. 3, 4. Mucin phenotypic parameters alone: The model based on MUC2 proportion, MUC6 proportion, incomplete Type II and incomplete Type III proportion achieved an AUC of 0.774 ± 0.043. Clinical parameters alone: The optimal model incorporating age, sex, PGII, HP status, and G17 achieved a carcinogenesis discrimination AUC of 0.764 ± 0.061. Integrated parameters: The combined model incorporating age, sex, PGII, HP status, G17, MUC2 proportion, MUC6 proportion,

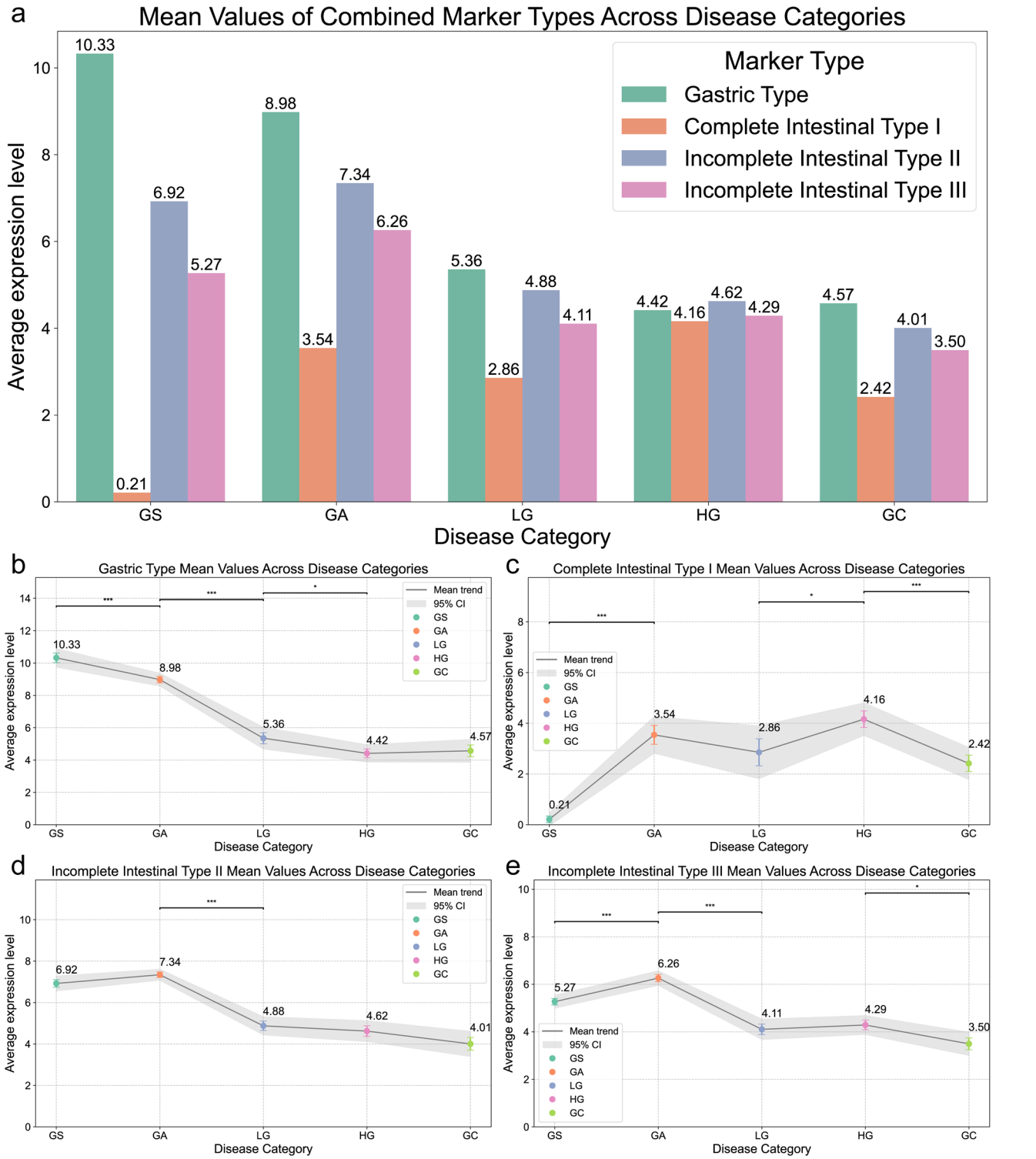

Fig. 3 | AUC in internal validation of four different mucin phenotypic markers. Classification characteristics of mucin phenotypes in different gastric diseases. **a** The expression of different mucin phenotype classifications in gastric diseases. **b**–**e** The trend of changes in gastric type, complete intestinal type I, incomplete intestinal type II, and type III in different gastric diseases. $^{*}P < 0.05$, $^{**}P < 0.01$; $^{***}P < 0.001$.

incomplete-type GIM type II proportion, and incomplete-type GIM type III proportion demonstrated superior performance (AUC = 0.814 ± 0.064), significantly outperforming single-parameter models (Fig. 8). These results confirm that incorporating mucin phenotypic parameters substantially improves predictive accuracy in stratifying malignant transformation in GIM.

Discussion

This study analyzed the expression patterns and phenotypic classification characteristics of mucin phenotypic markers (MUC5AC, MUC6, MUC2, and CD10) in different gastric diseases, and revealed the evolution of mucin phenotypic markers from GS to GA/GIM to IN to intestinal GC. And we further developed a MPMR model and its derivative, the MPMR-IMCP

Table 2 | Segmented tile statistics

Type	Train		Test		Validation		Total	
	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative
MUC5AC	37,509	78,892	8005	10,530	5375	17,534	50,889	10,6956
MUC6	25,351	89,690	4185	3336	2260	18,061	31,796	11,1087
MUC2	17,016	82,538	1767	10,749	2528	22,891	21,311	11,6178
CD10	5495	74,715	3267	11,736	1427	17,726	10,189	10,4177
Total	85,371	32,5835	17,224	36,351	11590	76,212	55,2583	

The number of tiles was calculated based on a size of 224 × 224 pixels. The Train and Test sets were derived from the FHCMU-ESD internal queue, while the Validation set was obtained from the ZH-EGD independent verification queue.

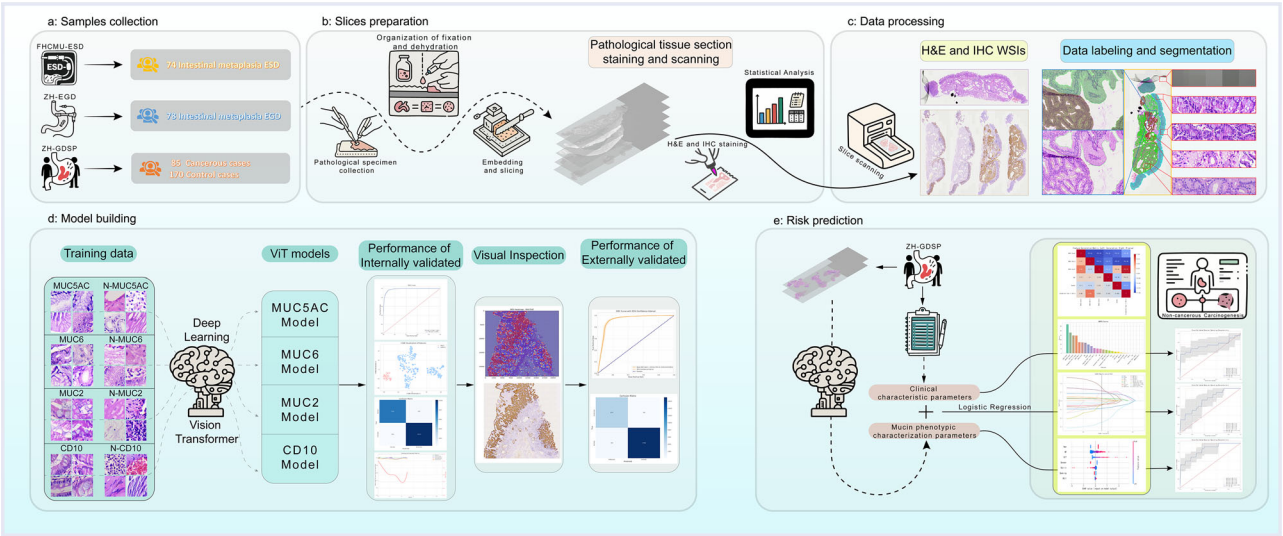

Fig. 4 | Flowchart system. The flowchart system illustrates the entire process of this study, from data collection to data processing, model development, and finally to risk assessment, clearly presenting the core stages through four color-coded sections: **a** (Data Collection) includes cohorts for training and testing datasets of deep learning models, cohorts for external validation of deep learning models, and cohorts for constructing cancer risk prediction models. **b** (Slices preparation) provides an overview of the preparation process of pathological tissue sections, including the preparation of pathological tissue sections, H&E and IHC staining, and statistical analysis of staining results. **c** (Data Processing) provides an overview of the pre-processing workflow for pathological tissue sections, including the scanning of sections, WSIs preparation, and the high-precision annotation process of H&E images based on IHC results. **d** (Model Construction) employs the ViT deep learning framework, where processed image data are input to train recognition models for four mucin phenotype markers (MUC5AC, MUC6, MUC2, and CD10), followed by efficacy validation on both internal and external datasets, with the generation of visual simulation heatmaps. **e** (Risk Prediction) incorporates clinical information from populations with a high incidence of GC, extracting clinical feature parameters suitable for modeling. Based on the predictive performance of deep learning models applied to H&E images, it derives mucin phenotype marker characteristics from high-incidence GC cases. Utilizing various feature selection methods, it constructs a carcinogenic risk prediction model for GIM, either independently or in combination, and performs comparative analyses to evaluate predictive efficacy.

model, through the integration of digital pathology and deep learning technologies. The MPMR model achieved, for the first time, high-precision prediction of four mucin markers (MUC5AC, MUC6, MUC2, and CD10) solely from H&E-stained images and introduced simulated staining heat maps as diagnostic visualization tools, while the MPMR-IMCP model significantly improved prediction of malignant transformation in GIM by fusing pathological imaging features with clinical data. This work overcomes the limitations of traditional special staining-dependent mucin analysis by establishing a cost-effective, multi-marker prediction approach using H&E images.

Mucins, as core functional molecules of the gastric mucosal barrier, exhibit dynamic expression patterns that not only reflect mucosal injury and repair status[35,36], but also actively participate in shaping the precancerous microenvironment[37]. Our study reveals dual characteristics of mucin phenotypes during gastric lesion progression. On one hand, gastric-type markers MUC5AC and MUC6 demonstrate progressive reduction (GS → GA → LGIN → HGIN), directly correlating with the destruction of gastric pit epithelium and glandular architecture, as well as loss of secretory function. Specifically, decreased MUC5AC reflects impaired pit epithelium and mucosal barrier function[38], while MUC6 reduction marks irreversible glandular atrophy[39]. These findings corroborate previous reports[15,40] linking MUC5AC loss in intestinal-type GC with poor tumor differentiation, collectively establishing gastric mucin depletion as a critical marker of malignant transformation. On the other hand, intestinal-type markers MUC2 and CD10 exhibit a characteristic nonlinear "rise-and-fall" pattern: significant increase during GS → GA, initial decline from GA → LGIN, and sharp reduction at HGIN → GC stages. This dynamic trajectory elucidates the transformation from GIM to carcinogenesis - the initial elevation represents adaptive repair response through incomplete GIM (types II/III), while simultaneously establishing a precancerous microenvironment[41]; whereas the subsequent abrupt decrease indicates tumor cells acquiring dedifferentiation capacity and losing mature intestinal epithelial phenotype[15,42]. Notably, this nonlinear progression highlights the limitations of single-timepoint GIM marker assessment, emphasizing the need to incorporate gastric mucin loss trends for comprehensive evaluation of GIM malignant potential.

Based on these dynamic mucin expression patterns, mucin phenotyping provides dual value for disease classification and risk stratification.

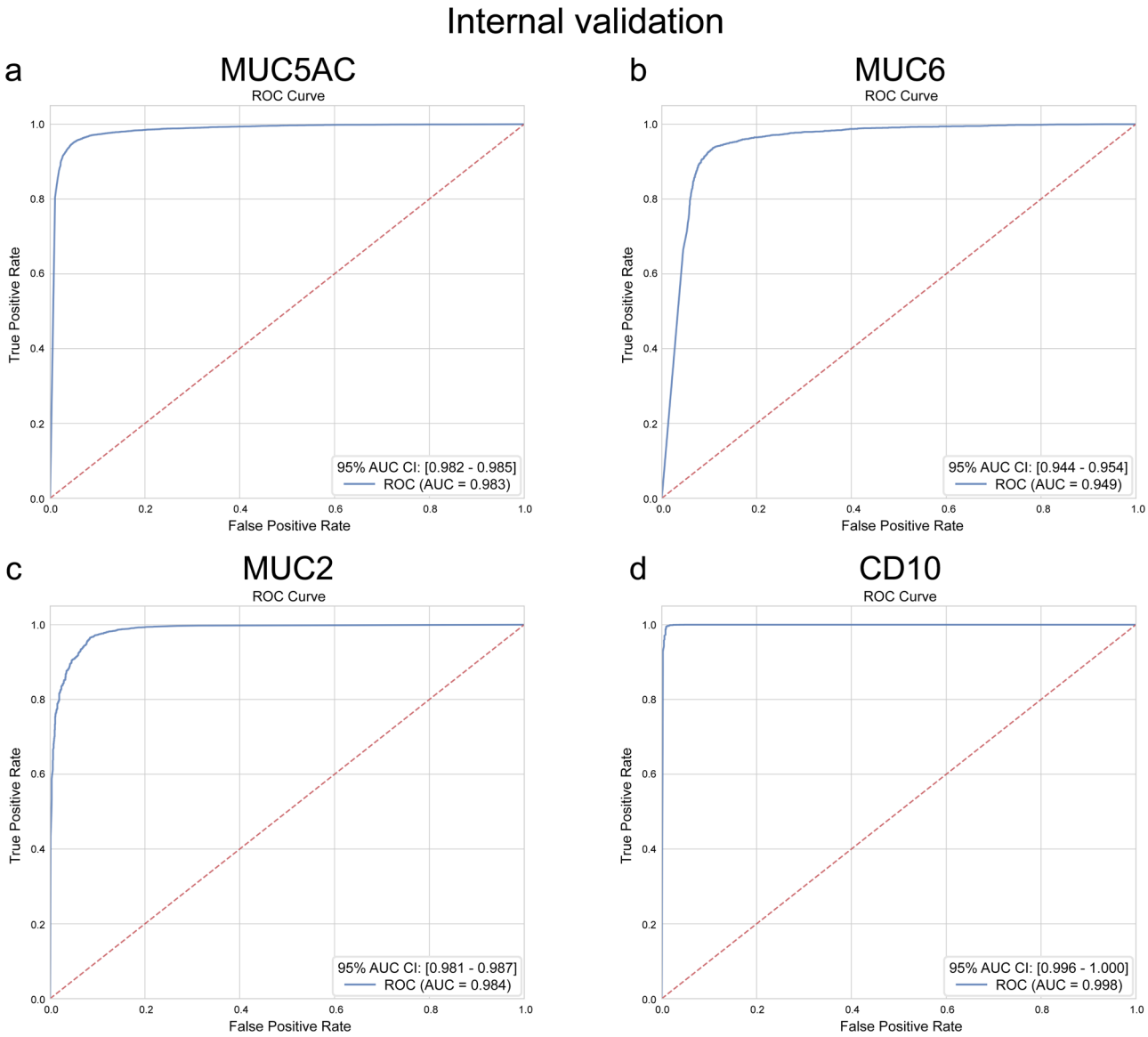

Fig. 5 | AUC in internal validation of four different mucin phenotypic markers. The predictive performance of MUC5AC, MUC6, MUC2, and CD10 in the FHCMU-ESD cohort was evaluated by plotting the ROC curves and calculating the AUC values of the model using the internal validation dataset, the model demonstrated strong AUC performance across all four markers. **a** MUC5AC. **b** MUC6. **c** MUC2. **d** CD10.

The combined expression profiles of gastric-type (MUC5AC + / MUC6 +) and intestinal-type (MUC2 + /CD10 +) markers accurately distinguish complete GIM (type I) from incomplete GIM (types II/III)[43–45]. This classification shows stage-specific correlations: coexistence of gastric-type and incomplete GIM (II/III) in GS/GA stages indicates early repair mechanism imbalance; progression to LGIN → HGIN shows gradual reduction of gastric-type and incomplete GIM with relative stability of complete GIM (I), marking the critical threshold of differentiation disorder; finally, in GC stage, complete disappearance of gastric-type is replaced by heterogeneous coexistence of complete GIM (I) and incomplete GIM (III), with type III showing significant association with highly malignant features[46,47]. These findings carry clear clinical implications that GA patients predominantly exhibiting incomplete GIM (III) require enhanced endoscopic surveillance, while GC patients retaining complete GIM (I) may demonstrate relatively favorable prognosis. Therefore, an integrated analysis strategy incorporating both dynamic mucin phenotype changes and subtype characteristics will provide novel perspectives for precision management of gastric precancerous lesions.

Traditional classification methods relying on manual interpretation of special stains suffer from low efficiency and substantial subjectivity. Existing artificial intelligence models primarily focus on tumor diagnosis and prognosis[48–52], leaving a critical gap in dedicated tools for precancerous lesion assessment. This study leverages a Vision Transformer (ViT) architecture integrated with UNI pretrained weights to extract complex pathological features under limited training samples, successfully establishing the MPMR model capable of identifying four mucin phenotypic markers. The interpretability of the model is further enhanced through visualized output of simulated staining heatmaps. It should be noted that under identical experimental settings, models initialized with UNI v1 weights achieved a mean AUC of 0.93 on internal validation, whereas switching to UNI v2 yielded significant performance improvement with a mean AUC of 0.98. External validation further confirmed the robust predictive performance of the UNI v2-based trained model ($AUC_{MUC5AC} = 0.994$, $AUC_{MUC6} = 0.997$, $AUC_{MUC2} = 0.980$, $AUC_{CD10} = 0.921$). As an objective quantitative tool, MPMR provides three distinct clinical advantages: First, it establishes a reproducible and standardized benchmark that mitigates inter- and intra-observer variability. Second, it integrates WSI-level contextual information

Table 3 | MPMR performance statistics

Type	AUC	Loss	Accuracy	Sensitivity	Specificity	F1 score	MCC
Model internal validation							
MUC5AC	0.988	0.033	0.958	0.962	0.964		0.896
Positive				0.932	0.949	0.940	
Negative				0.962	0.949	0.955	
MUC6	0.969	0.007	0.911	0.973	0.939		0.820
Positive				0.916	0.924	0.920	
Negative				0.905	0.895	0.900	
MUC2	0.990	0.023	0.976	0.983	0.981		0.606
Positive				0.466	0.898	0.632	
Negative				0.815	0.996	0.984	
CD10	0.998	0.018	0.997	0.999	0.993		0.976
Positive				0.985	0.971	0.978	
Negative				0.997	0.998	0.998	
Model external validation							
MUC5AC	0.994		0.976	0.975	0.980	0.984	0.935
MUC6	0.997		0.980	0.979	0.983	0.988	0.906
MUC2	0.980		0.915	0.910	0.961	0.951	0.684
CD10	0.921		0.949	0.967	0.723	0.972	0.652

The performance of MPMR was evaluated using multiple indicators, including AUC, loss value, accuracy, sensitivity, specificity, F1 score, and Matthews correlation coefficient. The internal validation performance of the model also includes statistics for positive and negative class samples.

to quantify subtle morphological features beyond human visual perception, thereby facilitating early detection of incipient or atypical lesions. Third, it serves as an efficient pre-screening and prioritization system that automatically flags cases warranting expert review, thereby optimizing diagnostic workflows and ensuring optimal allocation of specialist resources to the most complex diagnostic decisions. However, it should be particularly noted that the demonstrated "high accuracy" of MPMR reflects its concordance with histopathological gold standards. Therefore, the present performance should be interpreted as the model's strong potential in replicating diagnostic benchmarks, rather than claiming superiority over the comprehensive diagnostic capabilities of pathologist.

After successfully constructing the MPMR model in this study, we further utilized the long-term follow-up cohort of high-risk population in the high-incidence areas of GC in northern China to establish MPMR-IMCP by combining the pathological features derived from mucin phenotypes (MUC2/MUC6 and incomplete type II/III ratio) and integrating clinical parameters (age and multiple serological indices). MPMR-IMCP achieved an AUC of 0.814, representing a 6.5% improvement (ΔAUC = 0.05 over clinical-only models. This improvement stemmed from two key design strategies. First, a multi-scale feature fusion technique linking local mucin expression patterns (e.g., spatial distribution of incomplete type III) with global pathological features. Second, a dynamic weighting algorithm that balanced contributions from pathomics and clinical data to avoid feature redundancy. Compared with the traditional OLGIM system, MPMR-IMCP introduces quantitative mucin phenotype parameters for the first time, enhancing risk stratification resolution. Relative to multi-omics models[53,54], MPMR-IMCP directly extracts biomarker features from pathological images, eliminating the need for costly molecular testing. During the cross-validation process of IMCP, we observed performance variability across different folds, with image-only models outperforming the integrated model in specific subsets. We consider that the phenomenon can be attributed to two primary factors. First, certain folds contained a higher proportion of diagnostically challenging cases—particularly those with atypical morphological features or complex clinicopathological correlations—where relying on robust image features alone proved more effective than integrating potentially discordant multimodal data. Second, this observation highlights the inherent difficulty in achieving optimal fusion of heterogeneous data sources, particularly when dealing with rare but clinically significant edge cases. It should be specifically noted that despite these variations in individual folds, the multimodal integrated model maintained statistically superior average performance across all validation sets. This demonstrates that while unimodal approaches may occasionally excel in specific scenarios, the integrated framework provides more consistent and reliable performance across diverse clinical presentations. The observed variability actually reinforces the model's clinical relevance, as it reflects real-world diagnostic challenges where certain cases naturally favor either morphological or clinical analysis. These findings underscore the importance of maintaining multimodal integration while suggesting opportunities for future refinement through weighted fusion strategies or uncertainty-aware modeling. In short, MPMR-IMCP offers three major advantages: Biological interpretability: parameters such as MUC2/MUC6 and Incomplete Type II/III proportions directly reflect the molecular characteristics underlying metaplastic subtypes; Technical accessibility: requires only routine H&E slides without the need for specialized equipment; Dynamic prediction capability: enables real-time risk score updates through follow-up imaging. These innovations make MPMR-IMCP the first mucin phenotype-based tool for predicting malignant transformation in GIM, demonstrating high precision, cost-effectiveness, and dynamic adaptability.

This study presents a novel computational pathology framework (MPMR-IMCP) that predicts the GIM progression by integrating conventional H&E-stained images with clinical information. Adopting a two-stage architecture, the MPMR-IMCP first generates artificial biomarkers directly from H&E images, then fuses them with clinical parameters through a machine learning risk model. This approach overcomes the technical limitations of traditional staining-dependent methods and addresses crucial unmet needs in GIM risk stratification. The framework's innovation is characterized by three key aspects: the novel utilization of mucin phenotypic markers, effective integration of pathological features with clinical information, and enhanced model interpretability through simulated staining heatmaps. These features collectively ensure both technical rigor and significant clinical translational value of the proposed model. Notwithstanding these contributions, several limitations warrant consideration. Regarding clinical

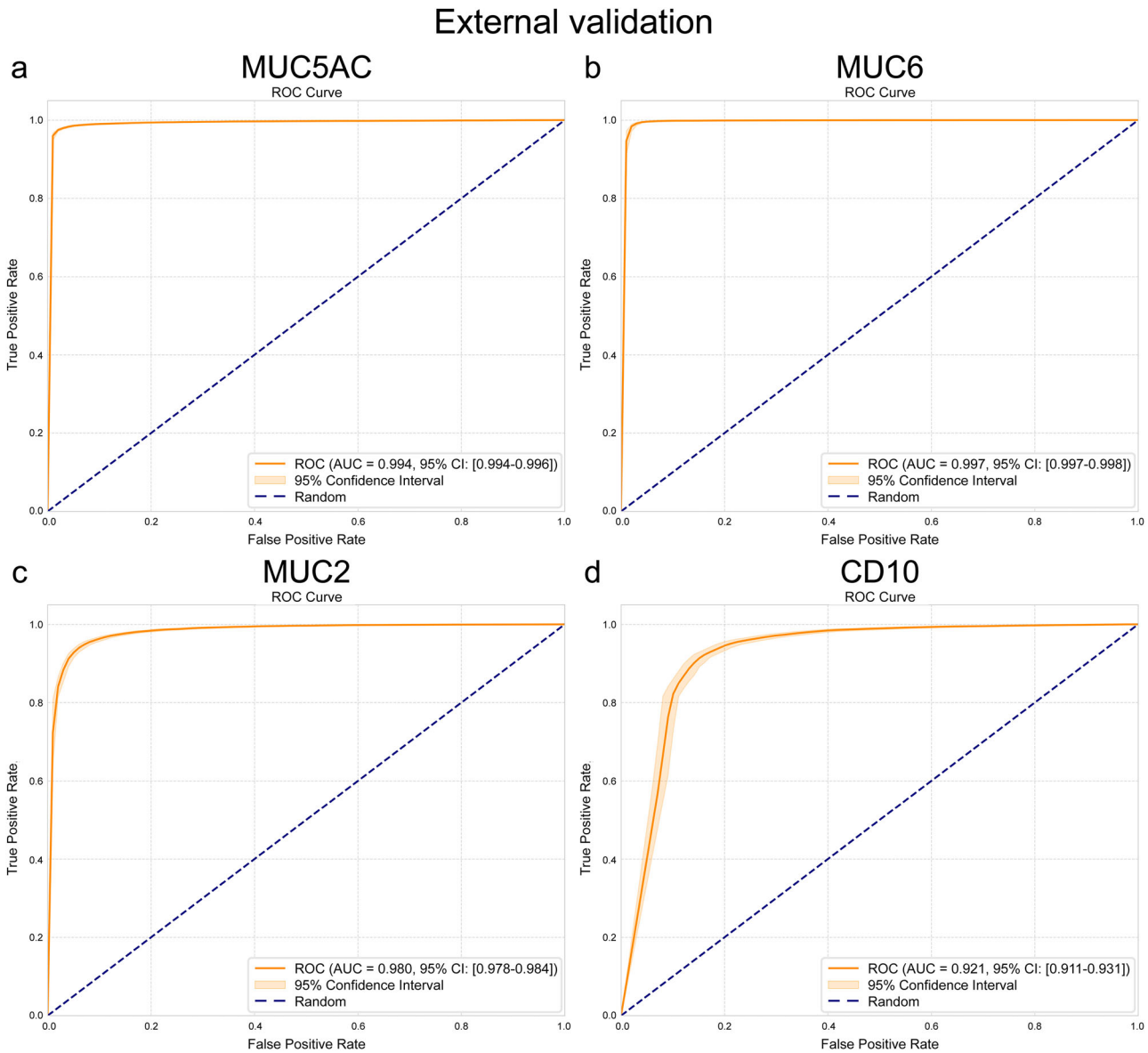

Fig. 6 | AUC in internal validation of four different mucin phenotypic markers. The predictive performance of MUC5AC, MUC6, MUC2, and CD10 in the ZH-EGD cohort was evaluated by plotting receiver operating characteristic (ROC) curves and calculating the AUC values of the model using an independent external validation dataset, the model maintained strong AUC performance in the external validation dataset. **a** MUC5AC. **b** MUC6. **c** MUC2. **d** CD10.

validation and applicability, the current study lacks head-to-head comparison with widely adopted endoscopic gastric cancer risk assessment systems such as the Kyoto classification[55]. While our model provides molecular insights beyond endoscopic capabilities, its incremental clinical value relative to established endoscopic standards requires further validation in cohorts integrating endoscopic and pathological data. Additionally, the simplification of Helicobacter pylori infection status to a binary variable (presence/absence) fails to distinguish between active infection and post-eradication states. The fundamentally distinct mucosal inflammatory microenvironments and tissue repair processes characterizing these phases may differentially influence mucin expression patterns and subsequent carcinogenesis risk, necessitating refined H. pylori status classification to enhance model precision. From methodological perspectives, the image augmentation strategies employed in modeling (including random grayscale conversion and color jitter) may distort the natural distribution of H&E stains, potentially compromising biological plausibility. Although the current model demonstrates robust performance, systematic evaluation of these augmentation methods' impact on pathological feature interpretation requires future investigation. Furthermore, multi-task learning—a more efficient and biologically interpretable paradigm enabling cross-marker feature sharing while preserving biomarker-specific signals—was not implemented due to considerations of task complexity, performance metrics, and sample size limitations. Subsequent research will prioritize exploration in this direction. Concerning data resources and generalizability, model training relied exclusively on pixel-level annotated data from Chinese high-risk populations, with relatively limited samples for clinical validation. This may impair model performance across diverse ethnic and geographical populations, necessitating validation of long-term efficacy through larger, more diverse multi-center cohorts. Moreover, while pixel-level manual annotation enhances training sample quality, its substantial costs and potential consistency issues in atypical or rare cases present challenges. Future investigations should examine model performance under varied supervised learning conditions. In summary, these limitations delineate clear directions for improvement. Prospective clinical validation, algorithmic architecture optimization, and multi-center data integration will further enhance the clinical applicability and scientific value of the MPMR-IMCP framework.

In summary, this study found that mucin phenotypic markers and their classification can be used as important reference indicators

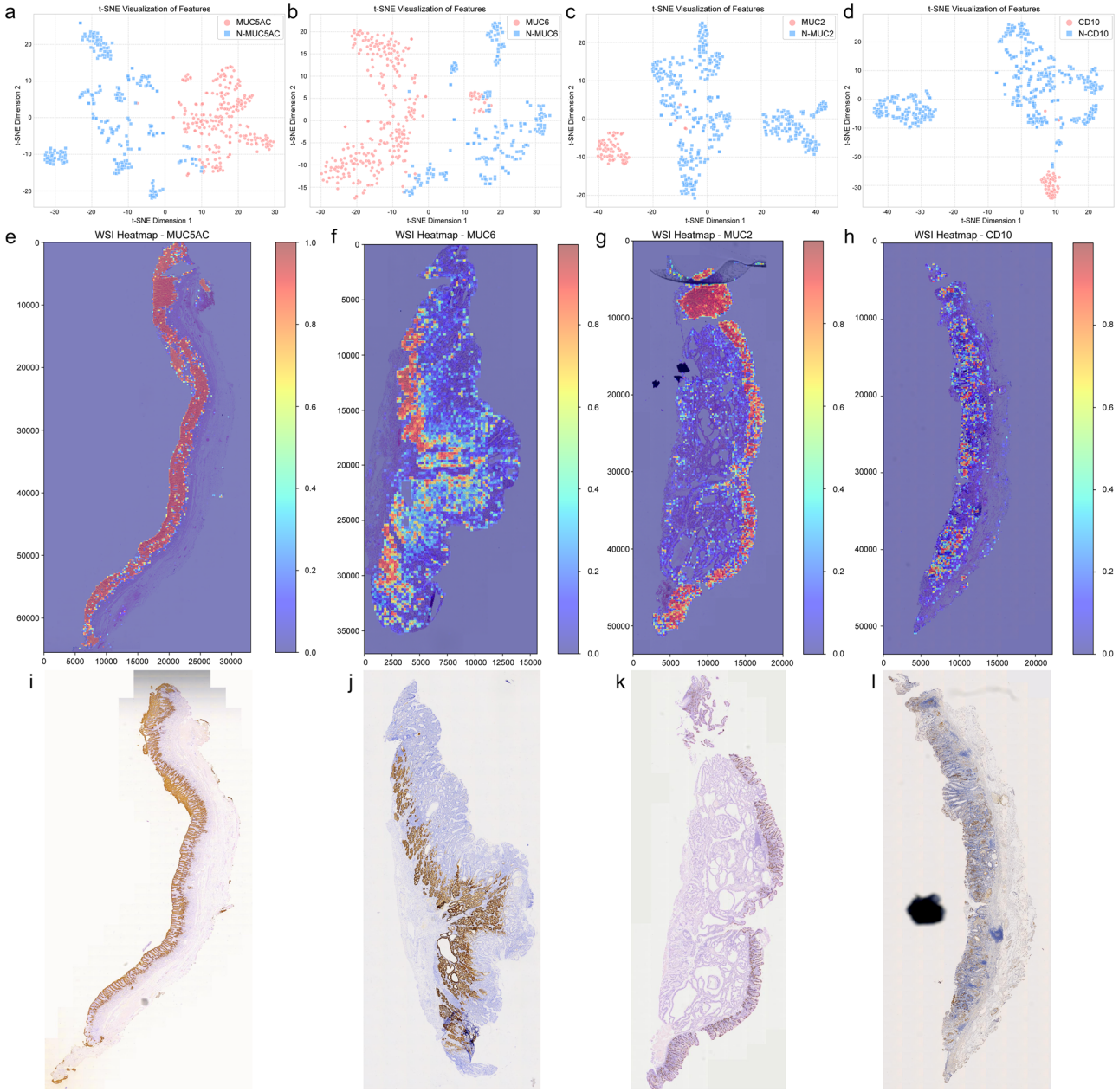

Fig. 7 | t-SNE feature space visualization and Heatmap of simulated staining. By applying the t-SNE feature space visualization method, the feature vector distributions of four different mucin phenotypic markers during the training process were plotted. t-SNE generates a reduced feature space where similar samples are represented by nearby points and different samples are represented by distant points. The results show that the vector features of the four labels are clearly separated during the training process, indicating that the model effectively learns to distinguish between positive and negative image features. Pathological tissue images for H&E staining were then predicted using MPMR, and simulated staining heat maps were generated and compared with IHC staining. **a**–**d** t-SNE feature space visualization. **e**–**h** Model-generated simulated staining heat maps predict the expression of MUC5AC, MUC6, MUC2, and CD10. **i**–**l** IHC staining of MUC5AC, MUC6, MUC2, and CD10.

Table 4 | MPMR-IMCP performance statistics

Type	AUC						Accuracy	Precision	Recall	F1 score	Log loss
	Mean	Fold 0	Fold 1	Fold 2	Fold 3	Fold 4					
Wsi	0.774	0.791	0.713	0.744	0.820	0.802	0.779	0.636	0.609	0.622	0.475
Clinical	0.764	0.861	0.812	0.713	0.715	0.721	0.753	0.591	0.565	0.578	0.484
Integrated	0.814	0.875	0.847	0.752	0.866	0.731	0.818	0.714	0.652	0.682	0.413

The performance of MPMR-IMCP was evaluated using multiple metrics, including AUC, accuracy, precision, recall, F1 score, and logistic regression loss. The predictive performance was assessed separately for models using only clinical feature parameters, only WSIs image feature parameters, and a combination of both types. Only the best-performing results under each parameter setting were recorded.

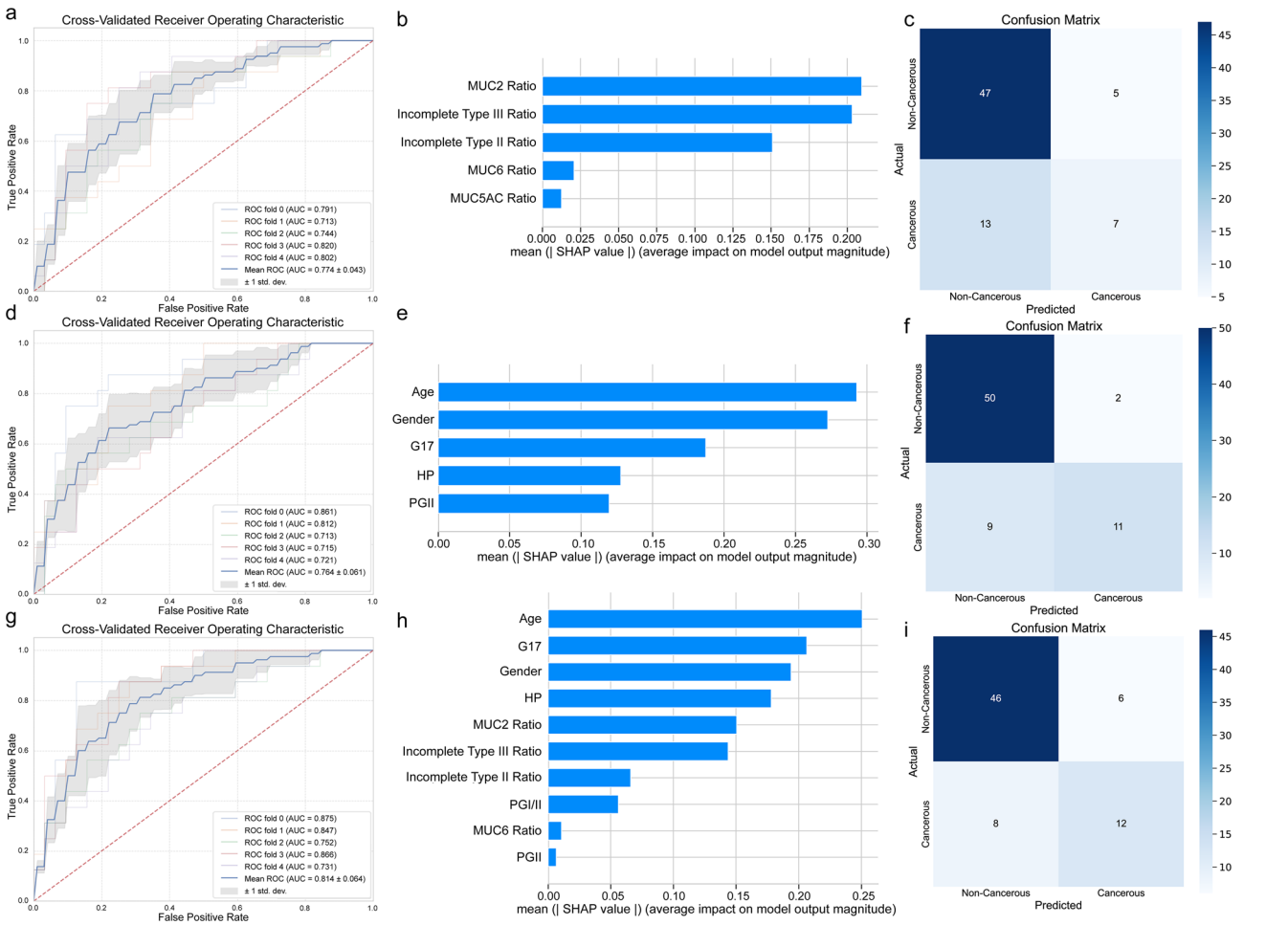

Fig. 8 | Prediction of malignant transformation in GIM. The malignant transformation prediction model for GIM was constructed using pathological tissue image characteristic parameters, clinical characteristic parameters, and a combination of both types. ROC curves, comprehensive Importance and confusion matrices were generated to evaluate model performance. **a–c** Efficacy of pathological tissue image characteristic parameters alone. **d–f** Efficacy of clinical characteristic parameters alone. **g–i** Efficacy of the combined model incorporating both types of characteristic parameters. The predictive performance of clinical and pathological image parameters was comparable when used independently, but the combined model demonstrated significantly improved efficacy in predicting malignant transformation in GIM.

to evaluate the lesion stage and gastric disease progression, and revealed that the gradual loss of gastric mucin (MUC5AC/MUC6) and the dynamic fluctuation of GIM-related markers (MUC2/CD10) in the precancerous stage and the decrease in the carcinogenesis stage during the evolution spectrum of gastric diseases. And successfully developed the first H&E image-based dual-functional model that integrates mucin phenotype recognition with prediction of malignant transformation in GIM. The MPMR model achieved high-accuracy prediction of four mucin markers using deep learning, with strong external validation performance (mean AUC = 0.973), laying a solid foundation for clinical implementation. The MPMR-IMCP model further improved the prediction of malignant transformation in GIM, reaching an AUC of 0.814 through the integration of pathomics and clinical data, and exhibited significant advantages over conventional methods. Together, these model fills crucial gaps in GIM risk assessment by enabling multi-marker prediction, pathology-clinical feature integration, and IHC-alternative virtual staining. Our findings offer actionable technical solutions for efficient recognition in GIM malignant transformation.

Methods

Study subjects

The FHCMU-IHC cohort for IHC staining analysis, which included both surgical and ESD specimens, comprised patients who underwent endoscopy and pathological tissue sampling and slide preparation at the First Hospital of China Medical University between October 2019 and December 2024. Except for GS, the other cases were accompanied by different degrees of pathological changes of GIM. Inclusion criteria included: (1) Age between 30 and 80 years; (2) Patients with pathological diagnosis of GS, GA, LGIN, HGIN and GC; (3) Histopathological changes of GIM; (4) Pathological tissue sections with H&E staining; (5) IHC stained sections with four mucin phenotypic markers (MUC5AC/MUC6/MUC2/CD10). Exclusion criteria included: (1) Patients with incomplete pathological diagnosis or clinical information; (2) The quality of slide preparation or scanning is unqualified (such as damaged, missing, wrinkled, stained tissue, out-of-focus, blurry, overlapping, or other obvious interference factors in the scan). The FHCMU-ESD cohort, which consisted exclusively of ESD specimens and was used to develop the MPMR internal dataset, comprised patients who underwent endoscopy and pathological tissue sampling and slide preparation at the First Hospital of China Medical University, between October 2019 and December 2021. Inclusion criteria were: (1) Pathologically confirmed GIM; (2) Availability of H&E-stained and IHC stained with four mucin phenotypic markers slides. Exclusion criteria were: (1) Incomplete clinical or pathological records; (2) Substandard slide preparation or scanning quality, such as tissue damage, missing sections, folds, stains, or scanning defects including out-of-focus areas, blurring, overlapping, or other significant artifacts. The ZH-EGD cohort, used for independent external validation of MPMR, included patients from the ZH-GDSP in Liaoning Province between February and December

1997[31,32], and followed the same inclusion and exclusion criteria as the FHCMU-ESD cohort and was composed of gastric biopsy specimens. The ZH-GDSP, which was used for MPMR-IMCP evaluation and consists of gastric biopsy specimens, included patients who completed gastric cancer screening and follow-up in the Zhuanghe High-Risk Area between February 1997 and June 2021[31,32]. Inclusion criteria were: (1) Pathological diagnoses of GS, gastric mucosal erosion or ulceration, GA, or GD; (2) Histopathological evidence of GIM; (3) Availability of H&E-stained slides; (4) Complete pathological diagnoses, serological indices, and follow-up data. Exclusion criteria were: (1) Incomplete pathological, serological, or follow-up records; (2) Cases of cancer or death within half a year after screening; (3) Substandard slide preparation or scanning quality, following the same criteria as applied to the FHCMU-ESD cohort. H&E and IHC slides used in all cohorts of the study were obtained from the same tissue block. The study protocol was approved by the Ethics Committee of the First Hospital of China Medical University (Ethical Approval Number: [2023]2023-446-2). All procedures were conducted in accordance with the Declaration of Helsinki. Written informed consent was signed by all participants.

Specimen quality assessment

To reflect real-world diagnostic practices, our study simultaneously included clinical surgical specimens, endoscopic biopsy specimens, and ESD specimens. All histopathological evaluations, including the assessment of GIM, were performed according to standardized international guidelines (WHO classification). This ensured consistent application of diagnostic criteria—such as the identification of dysplastic glands and nuclear features—across all specimen types, regardless of their origin. To address potential variability introduced by specimen size, feature extraction was conducted from multiple representative regions within each specimen, using standardized microscopic fields as the unit of assessment. Furthermore, in the subsequent multi-source risk prediction analysis, ' specimen type' and 'tissue area' were included as covariates to statistically control for potential size-related effects, thereby ensuring the robustness and comparability of our research findings across various clinical samples.

Immunohistochemical staining

The collected tissue samples were paraffin-embedded, sectioned, and dried. The paraffin sections underwent conventional dewaxing and hydration using xylene and gradient-free ethanol. Antigen retrieval was performed using EDTA under high-pressure heat treatment, followed by incubation with MUC5AC/MUC6/MUC2/CD10 antibodies. Antibodies used include: MUC5AC (Abcam, ab77576, 1:100 dilution), MUC6 (Abcam, ab216017, 1:250 dilution), MUC2 (Abcam, ab11197, 1:5000 dilution), CD10 (Abcam, ab270722, 1:500 dilution). After washing, secondary antibodies were applied for incubation. Finally, color development was achieved using freshly prepared DAB solution, counterstained with H&E, dehydrated, and permanently mounted. The sections used for IHC preparation were all derived from the same tissue blocks as the H&E sections.

Whole-slide image preparation

This study utilized conventional paraffin-embedded HE stained sections. The initial screening of the pathological sections was performed by pathologists using a low-power microscope (10× magnification, 1.04 μm/pixel) to exclude specimens with tissue defects, staining abnormalities, or vague glandular structures. In addition, the integrity of the glass slides was evaluated to eliminate scratches, breakages, or contaminants that could interfere with optical observation. Qualified specimens underwent surface cleaning and were subjected to whole-slide digital scanning at 40× magnification (0.26 μm/pixel) using the Motic EasyScan 1 digital slide scanner (China), generating digital pathological images in.mdsx format, which were archived on the server of the Department of Cancer Etiology and Screening Research at the First Hospital of China Medical University. Subsequently, image quality assessments were conducted using the Motic DSM professional viewing system. All procedures adhered strictly to clinical pathology technical standards.

Image annotation and preprocessing

The screened and converted image data were annotated with ROI by pathologists using the QuPath image annotation tool[33]. Using QuPath, pathologists marked staining regions corresponding to different mucin phenotypic markers (MUC5AC, MUC6, MUC2, and CD10) on H&E-stained images, referencing the associated special-stained slides. The OpenSlide library[56] was used to process ROI-annotated images and their corresponding annotation files, ensuring that segmented image patches retained their original parent labels. Images were divided into 224 × 224-pixel patches and stored in subfolders named according to WSIs case numbers to prevent information leakage, and were allocated across different training tasks, resulting in 552,583 labelled patches (detailed in Table 2). Image preprocessing was integrated into the model training workflow. Before inputting data into the model, the Torchvision Transform module[57] performed a series of image transformations and data augmentation steps, including: (1) removal of blank or contaminated images; (2) random horizontal and vertical flipping; (3) random rotation; (4) random affine transformation; (5) random color jittering; (6) random grayscale conversion; (7) standardization by normalizing the mean and standard deviation of RGB channels. Validation datasets underwent identical preprocessing procedures to ensure consistency and prevent validation bias.

MPMR model construction

During MPMR model development, we constructed a ViT model using pre-trained weights from the UNI-2 pathology foundation model. The core of the ViT architecture consists of multiple identical transformer encoder layers, each comprising two main components: multi-head attention and a multi-layer perceptron (MLP)[25]. UNI is a universal pathological self-supervised model pre-trained on over 100,000 diagnostic H&E-stained WSIs (more than 77 TB of data), encompassing 20 major tissue types and over 100 million images[27]. We employed its higher-performing v2 version as the pre-trained weights. The training parameters were set as follows: batch size = 64, initial learning rate = 0.0001, weight decay = 0.5, using the Adam optimizer and cross-entropy loss function. In terms of the fine-tuning strategy, we adopted a partial fine-tuning approach focused on efficiency and preventing overfitting: instead of performing full-parameter fine-tuning on the entire UNI v2 model, we primarily adjusted the query/key/value linear projection layers (qkv layers) in the attention mechanism, the fully connected layers (fc1 and fc2) in the MLP block, as well as the Dropout layer and the final linear classification head, while keeping the rest of the backbone network parameters frozen. This strategy allowed us to leverage the powerful pre-trained features of UNI v2 while efficiently adapting it to the target task and effectively controlling model complexity. The entire training was conducted with a fixed training-validation split (8:2), for a maximum of 100 epochs, and early stopping with a patience of 5 was applied. Data loading utilized 24 parallel workers to maximize efficiency. These configurations contributed to the optimization of training efficiency and stability. All subsequent analyses and test-set evaluations were conducted with the model selected as the best-performing on the independent internal validation set.

ZH-GDSP data processing

The study included 240 cases from the ZH-GDSP, collecting 11 clinical parameters: disease classification, age, sex, pathological diagnosis, five serological indices (PGI, PGII, PGI/II ratio, HP status, G17), smoking/alcohol history, and family history. Follow-up data through January 2025 identified 80 carcinogenesis cases (cancer group), which were matched 1:2 by age and sex with 160 controls. Categorical variables (age, sex, pathology, behavioral histories) were label-encoded, while serological indices were standardized using Z-score normalization.

MPMR-IMCP model construction

Multiple statistical and machine learning methods were used to evaluate feature relevance, including Pearson correlation, ANOVA, LASSO

regression, random forest feature importance, and SHAP values. Feature rankings were derived by integrating weighted scores from the ANOVA, LASSO, and random forest methods. Logistic regression was used as the baseline model, and sequential feature addition (top 1, 2, …, N features) was applied to identify optimal feature subsets based on performance evaluation. The selected best subset improved both prediction accuracy and model robustness.

Model visualization and interpretation

For MPMR, t-SNE was used to visualize feature vectors extracted from the final layers of the ViT model[34], while simulated heatmaps were compared with IHC gold standards. MPMR-IMCP employed SHAP (SHapley Additive exPlanations), with summary plots used to illustrate individual feature contributions.

Computational configuration

All analyses were conducted on a Linux CentOS 7.9 server using Python 3.10.15 and the PyTorch 2.2.2 deep learning framework (CUDA 12.4, cuDNN 8.9.2). Hardware specifications included: 4×NVIDIA GeForce RTX 3090 GPUs (24 GB GDDR6X memory each), 2×Intel Xeon Gold 6226 CPUs (12-core, 2.70 GHz), and 256 GB DDR4 RAM. Code development was carried out using Visual Studio Code v1.85 with Conda 4.12.0 for environment management.

Statistical analysis

MPMR model performance was evaluated using the AUC, accuracy, sensitivity (true positive rate), and specificity (true negative rate). MPMR-IMCP employed additional metrics, including precision (positive predictive value), recall (sensitivity), F1 score (harmonic mean of precision and recall), and calibration slope. All analyses were conducted using Python 3.11.5 libraries.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

Our code components, including image annotation, preprocessing, model training, and statistical analysis, can be reproduced using the methodological details provided in the 'Methods' section, which outlines data preprocessing and augmentation, model architecture, and training configurations. Image annotation was performed using QuPath (https://github.com/qupath/qupath), preprocessing was conducted with the OpenSlide library (https://github.com/openslide/openslide-python), the deep learning architecture was based on the ViT with UNI-2 pre-trained weights (https://github.com/mahmoodlab/UNI), and malignant transformation prediction models were built using Logistic Regression. We provide publicly available code and pretrained shared weights for access and download (https://github.com/ DowNxDawN/MPMR-IMCP).

Received: 30 July 2025; Accepted: 27 November 2025;
Published online: 12 December 2025

References

1. Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263 (2024).
2. Jing, J.-J. et al. Gastric cancer incidence and mortality in Zhuanghe, China, between 2005 and 2010. *World J. Gastroenterol.* **18**, 1262–1269 (2012).
3. Zivny, J., Wang, T. C., Yantiss, R., Kim, K. H. & Houghton, J. Role of therapy or monitoring in preventing progression to gastric cancer. *J. Clin. Gastroenterol.* **36**, S61–S62 (2003).
4. Correa, P. & Piazuelo, M. B. The gastric precancerous cascade. *J. Digestive Dis.* **13**, 2–9 (2012).
5. Lauren, P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma: an attempt at a histo-clinical classification. *Acta Pathol. Microbiol Scand.* **64**, 31–49 (1965).
6. Marcos, P., Gonçalves, N., Gama, J., Areia, M. & Dinis-Ribeiro, M. Standardizing endoscopic protocols increases detection of gastric cancer during surveillance of patients with precancerous conditions: a systematic review and meta-analysis. *Dig. Liver Dis.* **57**, 1588–1602 (2025).
7. Benites-Goñi, H. et al. OLGA and OLGIM staging systems on the risk assessment of gastric cancer: a systematic review and meta-analysis of prospective cohorts. *Ther. Adv. Gastroenterol.* **18**, 1756–2848 (2025).
8. Bogdanova, I. et al. Role of pre-existing incomplete intestinal metaplasia in gastric adenocarcinoma: A retrospective case series analysis. *World J. Clin. Cases* **11**, 2708–2715 (2023).
9. Breugelmans, T. et al. The role of mucins in gastrointestinal barrier function during health and disease. *Lancet Gastroenterol. Hepatol.* **7**, 455–471 (2022).
10. Jinfeng, M. et al. Expression of MUC5AC and MUC6 in invasive ductal carcinoma of the pancreas and relationship with prognosis. *Int J. Gastrointest. Canc* **34**, 9–18 (2003).
11. Park, J. S. et al. Immunohistochemical Expressions of MUC2, MUC5AC, and MUC6 in Normal, Helicobacter pylori Infected and Metaplastic Gastric Mucosa of Children and Adolescents. *Helicobacter* **20**, 260–268 (2015).
12. Dw, K. Mucins in cancer: function, prognosis and therapy. *Nat. Rev. Cancer* **9**, 874–885 (2009).
13. Shichijo, S., Hirata, Y., Yoshida, S., Yamada, A. & Koike, K. Tu1700 Endoscopic Findings for Risk Stratification of Gastric Cancer According to "Kyoto Classification of Endoscopic Gastritis. *Gastrointest. Endosc.* **81**, AB565 (2015).
14. Patel, N. M. et al. The Role of Mucin Expression in the Diagnosis of Oesophago-Gastric Cancer: A Systematic Literature Review. *Cancers (Basel)* **15**, 5252 (2023).
15. Akazawa, Y. et al. Clinicopathological and molecular characterization of early gastric adenocarcinoma in Helicobacter pylori-uninfected patients: emphasis on differentiated gastric adenocarcinoma. *J. Gastroenterol.* **57**, 725–734 (2022).
16. Gené, M. et al. Alterations in p53, Microsatellite Stability and Lack of MUC5AC Expression as Molecular Features of Colorectal Carcinoma Associated with Inflammatory Bowel Disease. *Int J. Mol. Sci.* **24**, 8655 (2023).
17. Zheng, X. et al. A deep learning model and human-machine fusion for prediction of EBV-associated gastric cancer from histopathology. *Nat. Commun.* **13**, 2790 (2022).
18. Shamai, G. et al. Deep learning-based image analysis predicts PD-L1 status from H&E-stained histopathology images in breast cancer. *Nat. Commun.* **13**, 6753 (2022).
19. Pati, P. et al. Accelerating histopathology workflows with generative AI-based virtually multiplexed tumour profiling. *Nat. Mach. Intell.* **6**, 1077–1093 (2024).
20. Anand, D. et al. Deep Learning to Estimate Human Epidermal Growth Factor Receptor 2 Status from Hematoxylin and Eosin-Stained Breast Tissue Images. *J. Pathol. Inf.* **11**, 19 (2020).
21. Liu, Y. et al. Predict Ki-67 Positive Cells in H&E-Stained Images Using Deep Learning Independently From IHC-Stained Images. *Front Mol. Biosci.* **7**, 183 (2020).
22. Jin, D. et al. Teacher-student collaborated multiple instance learning for pan-cancer PDL1 expression prediction from histopathology slides. *Nat. Commun.* **15**, 3063 (2024).
23. Han, K. et al. A Survey on Vision Transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 87–110 (2023).
24. Xie, H. et al. Deep-Learning-Based Few-Angle Cardiac SPECT Reconstruction Using Transformer. *IEEE Trans. Radiat. Plasma Med Sci.* **7**, 33–40 (2023).

25. Dosovitskiy, A. et al. An image is worth 16x16 words: transformers for image recognition at scale. In *Proceedings of the 9th International Conference on Learning Representations (ICLR),* 611−631 (2021).
26. Lu, M. Y. et al. A visual-language foundation model for computational pathology. *Nat. Med* **30**, 863–874 (2024).
27. Chen, R. J. et al. Towards a general-purpose foundation model for computational pathology. *Nat. Med* **30**, 850–862 (2024).
28. Ding, T. et al. A multimodal whole-slide foundation model for pathology. *Nat. Med* **31**, 3749–3761 (2025).
29. Verma, J. et al. From slides to insights: Harnessing deep learning for prognostic survival prediction in human colorectal cancer histology. *Open Life Sci.* **18**, 20220777 (2023).
30. Luo, J., Pan, M., Mo, K., Mao, Y. & Zou, D. Emerging role of artificial intelligence in diagnosis, classification and clinical management of glioma. *Semin Cancer Biol.* **91**, 110–123 (2023).
31. Tu, H. et al. A Serological Biopsy Using Five Stomach-Specific Circulating Biomarkers for Gastric Cancer Risk Assessment: A Multi-Phase Study. *Am. J. Gastroenterol.* **112**, 704–715 (2017).
32. Sun, L., Wen, Z., Zhang, Y. C. & Yuan, Y. From the Gastric Cancer High Incidence Area to Healthy County National Demonstration Base of Health Management: A 40-Year Review of Gastric Cancer Prevention and Control Program in Zhuanghe City of Liaoning Province. *China Cancer* **33**, 790–798 (2024).
33. Bankhead, P. et al. QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
34. Van Der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
35. Babu, S. D., Jayanthi, V., Devaraj, N., Reis, C. A. & Devaraj, H. Expression profile of mucins (MUC2, MUC5AC and MUC6) in Helicobacter pylori infected pre-neoplastic and neoplastic human gastric epithelium. *Mol. Cancer* **5**, 10 (2006).
36. Kang, H. M. et al. Effects of Helicobacter pylori Infection on gastric mucin expression. *J. Clin. Gastroenterol.* **42**, 29–35 (2008).
37. Oosterlinck, B. et al. Mucin-microbiome signatures shape the tumor microenvironment in gastric cancer. *Microbiome* **11**, 86 (2023).
38. Vernygorodskyi, S. Immunohistochemical evaluation of mucin expression in precancerous tissue of stomach. *Exp. Oncol.* **35**, 114–117 (2013).
39. Wang, Y. et al. Histopathological features of glandular atrophy of the lamina propria of the gastric mucosa during its occurrence and development. *BMC Gastroenterol.* **23**, 395 (2023).
40. Miao, L. et al. Unique immunohistochemical profiles of MUC5AC, MUC6, P53, and Ki67 in gastric atypical hyperplasia and dysplasia. *Int J. Clin. Exp. Pathol.* **17**, 63–71 (2024).
41. Jass, J. R., Strudley, I. & Faludy, J. Histochemistry of epithelial metaplasia and dysplasia in human stomach and colorectum. *Scand. J. Gastroenterol. Suppl.* **104**, 109–130 (1984).
42. Vilarinho, T., Pádua, D., Pereira, B., Mesquita, P. & Almeida, R. MISP Is Overexpressed in Intestinal Metaplasia and Gastric Cancer. *Curr. Oncol.* **31**, 2769–2779 (2024).
43. Mozgovoĭ, S., Livzan, M., Predvechnaia, I. & Kononov, A. Proliferative and mucin-producing activities in the foci of intestinal metaplasia associated with chronic atrophic gastritis and gastric ulcer. *Vestnik Rossiiskoi akademii meditsinskikh nauk.* **7**, 15–18 (2010).
44. Mabuchi, N. et al. Characteristics of gastric and intestinal mucin phenotypes of gastric carcinoma. *Hepatogastroenterology* **55**, 2277–2281 (2008).
45. Barresi, V., Vitarelli, E., Grosso, M., Tuccari, G. & Barresi, G. Relationship between immunoexpression of mucin peptide cores MUC1 and MUC2 and Lauren's histologic subtypes of gastric carcinomas. *Eur. J. Histochem* **50**, 301–309 (2006).
46. Rokutan, H. et al. Genomic and Pathologic Profiling of Very Well-Differentiated Gastric Adenocarcinoma of Intestinal Type: A Study With Emphasis on Diffuse-Type Transformation. *Am. J. Surg. Pathol.* **48**, 652–661 (2024).
47. Oya, M., Yao, T., Nakamura, T., Nakanishi, K. & Tsuneyoshi, M. Intestinal phenotypic expression of gastric depressed adenomas and the surrounding mucosa. *Gastric Cancer* **6**, 179–184 (2003).
48. Jin, L. et al. A multi-center performance assessment for automated histopathological classification and grading of glioma using whole slide images. *iScience* **26**, 108041 (2023).
49. Shi, Z. et al. A Two-Stage End-to-End Deep Learning Framework for Pathologic Examination in Skin Tumor Diagnosis. *Am. J. Pathol.* **193**, 769–777 (2023).
50. Sharma, A., Kumar, R. & Garg, P. Deep learning-based prediction model for diagnosing gastrointestinal diseases using endoscopy images. *Int. J. Med. Inform.* **177**, 105142 (2023).
51. Liu, L. et al. Diagnosis and segmentation effect of the ME-NBI-based deep learning model on gastric neoplasms in patients with suspected superficial lesions - a multicenter study. *Front Oncol.* **12**, 1075578 (2022).
52. Jiang, Y. et al. Biology-guided deep learning predicts prognosis and cancer immunotherapy response. *Nat. Commun.* **14**, 5135 (2023).
53. Huang, K. K. et al. Spatiotemporal genomic profiling of intestinal metaplasia reveals clonal dynamics of gastric cancer progression. *Cancer Cell* **41**, 2019–2037.e8 (2023).
54. Eskandarion, M. R. et al. Prediction of novel biomarkers for gastric intestinal metaplasia and gastric adenocarcinoma using bioinformatics analysis. *Heliyon* **10**, e30253 (2024).
55. Haruma, K., Kato, M., Inoue, K., Murakami, K. & Kamada, T. Kyoto classification of gastritis 3st ed. *Tokyo: Nihon Medical Center* (2023).
56. Goode, A., Gilbert, B., Harkes, J., Jukic, D. & Satyanarayanan, M. OpenSlide: A vendor-neutral software foundation for digital pathology. *J. Pathol. Inform.* **4**, 27 (2013).
57. TorchVision maintainers & contributors. TorchVision: PyTorch's Computer Vision library. *GitHub repository* GitHub (2016).

Acknowledgements

This study was supported by the National Key R&D Program of China (2022YFC2505100), Noncommunicable Chronic Diseases-National Science and Technology Major Project (2023ZD0501400), and The National Natural Science Foundation of China (82574197).

Author contributions

Y.Y. and H.T. contributed to the study design and revised the manuscript. X.W. completed the ROI annotation of pathological tissue images, the construction and training of models, the statistical analysis, and wrote the manuscript. F.W. provided pathological knowledge support and verification of ROI annotation. W.D. participated in the modeling analysis. C.N. completed the IHC staining of the external verification cohort. L.S. and Y.G. completed pathological diagnosis and cohort follow-up. N.D., Z.W., and L.L. performed IHC and WSI preparation and mucin staining. Q.X., J.J., and S.S. participated in the analysis of mucin expression. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41698-025-01224-w.

Correspondence and requests for materials should be addressed to Huakang Tu or Yuan Yuan.