British Journal of Hospital Medicine

Br. J. Hosp. Med. (Lond) **2026**; 87(9): 52019
https://doi.org/10.31083/BJHM52019

Article

Diagnostic Value of Serum HE4, TFF3, and CA72-4 for Gastric Cancer Stratified by *Helicobacter pylori* Infection Status

Muyong Wang[1,*], Mengke Xu[2]

[1]Department of General Surgery, Jinshazhou Hospital of Guangzhou University of Chinese Medicine, 510168 Guangzhou, Guangdong, China
[2]College of Health and Medicine, Dezhou University, 253023 Dezhou, Shandong, China
*Correspondence: mutongxianzi0318@163.com (Muyong Wang)
Academic Editor: John Alcolado
Submitted: 18 March 2026 Revised: 8 May 2026 Accepted: 27 May 2026 Published: 10 September 2026

Abstract

Aims/Background: Gastric cancer remains a leading cause of cancer-related mortality worldwide. *Helicobacter pylori* (Hp) infection is a major risk factor, but the diagnostic performance of serum biomarkers may vary according to Hp infection status. This study aimed to evaluate the diagnostic value of serum human epididymis protein 4 (HE4), trefoil factor 3 (TFF3), and carbohydrate antigen 72-4 (CA72-4), individually and in combination, for gastric cancer stratified by Hp infection status using a propensity score-matched retrospective case-control design. **Methods:** This retrospective study was conducted at Jinshazhou Hospital of Guangzhou University of Chinese Medicine, from January 2022 to December 2025. Propensity score matching (1:1 nearest neighbor matching, caliper 0.1) was performed to balance age, sex, body mass index (BMI), smoking history, alcohol consumption history, chronic kidney disease, chronic liver disease, and Hp infection status between groups. After matching, 268 gastric cancer patients and 268 healthy controls were included. Serum levels of HE4, TFF3, and CA72-4 were measured. Between-group comparisons of each marker, stratified by Hp status, were performed. Logistic regression was used to construct combined diagnostic models, and receiver operating characteristic (ROC) curve analysis was performed to evaluate the area under the curve (AUC). **Results:** After matching, all covariates achieved good balance with standardized mean differences <0.1. In the total population, serum levels of HE4, TFF3, and CA72-4 were significantly higher in the gastric cancer group than in the control group (all $p < 0.001$). Multivariate logistic regression identified all three markers as independent risk factors for gastric cancer (HE4: odds ratio [OR] = 1.05, 95% confidence interval [CI]: 1.03–1.06; TFF3: OR = 1.19, 95% CI: 1.13–1.26; CA72-4: OR = 1.46, 95% CI: 1.34–1.62; all $p < 0.001$). The combined model yielded an AUC of 0.899 (0.872–0.925) in the total population, significantly superior to any single marker (all $p < 0.001$). After stratification by Hp status, the AUC of TFF3 was significantly higher in Hp-positive individuals than in Hp-negative individuals ($p < 0.001$); no significant differences were observed for HE4 or CA72-4 between the two strata ($p > 0.05$). **Conclusion:** The diagnostic value of serum HE4, TFF3, and CA72-4 for gastric cancer varies according to Hp infection status. In Hp-positive patients, TFF3 demonstrates significantly higher diagnostic performance compared with Hp-negative individuals. HE4 and CA72-4 demonstrate stable performance across Hp strata, with no statistically significant differences observed. The combined detection of these biomarkers demonstrates favorable diagnostic performance regardless of Hp status. These findings suggest that Hp status-stratified testing provides an auxiliary strategy for optimizing differential diagnosis of gastric cancer.

Keywords: gastric cancer; *Helicobacter pylori*; HE4; TFF3; CA72-4

1. Introduction

Gastric cancer is one of the most common malignancies worldwide, with persistently high incidence and mortality rates. According to the 2020 global cancer statistics, there were over 1 million new cases of gastric cancer annually and approximately 769,000 deaths, ranking it the fifth most common cancer and the fourth leading cause of cancer-related death globally [1,2]. A recently published global burden study projected that among the 2008–2017 birth cohorts, 76% of the estimated 15.6 million lifetime gastric cancer cases could be attributed to *Helicobacter pylori* infection [3]. Despite advances in diagnostic and therapeutic techniques, the insidious onset of early-stage disease means that most patients are diagnosed at advanced stages, leading to a poor overall prognosis with 5-year survival rates below 30% [4]. Therefore, exploring effective early screening and diagnostic strategies is crucial for improving survival outcomes in gastric cancer patients.

Currently, gastroscopy combined with pathological biopsy remains the "gold standard" for gastric cancer diagnosis; however, its invasiveness, high cost, and limited acceptance restrict its application in large-scale population screening [5,6]. Serum tumor marker testing offers advantages, including non-invasiveness, simplicity, and good reproducibility, making it a potential auxiliary screening tool. However, commonly used clinical markers such as carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA19-9) demonstrate suboptimal sensitivity and specificity for early gastric cancer [7]. Therefore, identifying and validating novel, efficient serum markers or combinatorial models represents an important direction in early gastric cancer diagnosis research.

Classified as a Group I carcinogen by the International Agency for Research on Cancer, *Helicobacter pylori* (Hp) infection is the predominant risk factor for gastric cancer. Hp infection promotes gastric carcinogenesis by inducing chronic inflammation, gastric mucosal atrophy, intestinal metaplasia, and other precancerous lesions in a process known as Correa's cascade [8,9,10,11]. Notably, the population attributable fraction of Hp for non-cardia gastric cancer is as high as 71.2% in East Asia [10], and chronic atrophic gastritis with intestinal metaplasia represents a critical precancerous stage [11]. Furthermore, Hp infection itself may affect baseline levels of certain serum markers. A study has shown that Hp infection significantly upregulates the expression of pro-inflammatory factors such as interleukin (IL)-6 and IL-8 in the gastric mucosa [12], which in turn can regulate the synthesis of various acute-phase proteins. Additionally, intestinal metaplasia induced by Hp infection is closely associated with the expression of trefoil factor family proteins [13]. A recently published review systematically summarizes the application progress of Hp-related serum indicators in gastric cancer screening, emphasizing that Hp infection status is an important modifying factor affecting the diagnostic performance of serum markers [8]. These mechanistic insights suggest that Hp infection status may serve as an important effect modifier in serum marker-based gastric cancer diagnosis, yet this factor has not received adequate attention in current research.

In recent years, several novel serum protein markers have gained attention for their diagnostic value in gastric cancer. Human epididymis protein 4 (HE4), initially widely studied as an ovarian cancer marker, has subsequently been found to be significantly elevated in the serum of gastric cancer patients, potentially associated with tumor invasion and metastasis [14,15]. A recent study confirmed that HE4 expression intensity is significantly correlated with 5-year survival in gastric cancer patients, suggesting its role as a prognostic marker [15]. A previous *in vitro* study demonstrated that Hp infection may upregulate acute-phase protein expression through IL-6–dependent pathways; for instance, inter-alpha-trypsin inhibitor heavy chain H4 (ITIH4) has been confirmed to be regulated through this mechanism [16], suggesting that HE4 expression may similarly be influenced by Hp-mediated inflammatory responses. Trefoil factor 3 (TFF3) is a small-molecule protein primarily expressed in gastrointestinal mucosa, playing a role in mucosal protection and repair. Research has confirmed that Hp infection directly induces *TFF3* messenger RNA (mRNA) expression in gastric epithelial cells [13], and TFF3 levels are elevated in patients with chronic atrophic gastritis, with its diagnostic value varying according to Hp infection status [17,18]. A study in a high-risk population further demonstrated that serum TFF3 can serve as a non-invasive biomarker for gastric cancer [18]. Carbohydrate antigen 72-4 (CA72-4) is a classic gastric cancer-associated marker used in recurrence monitoring; however, its utility for early diagnosis is limited when used alone [19]. Notably, a multicenter study encompassing 7757 healthy adults demonstrated that Hp infection was significantly associated with elevated serum CA72-4 levels (odds ratio [OR] = 1.44), and pathological conditions such as atrophic gastritis also led to CA72-4 elevation [20]. These findings indicate that even this traditional marker may be influenced by Hp infection status and associated precancerous lesions.

Individual detection of these markers is limited by suboptimal performance, whereas combining multiple markers to construct predictive models may enhance diagnostic efficacy through complementary effects [21]. Multivariate analytical approaches have been advocated for integrating multiple biomarkers and clinical factors to improve gastric cancer risk assessment [21]. However, no studies to date have systematically evaluated the diagnostic value of serum HE4, TFF3, and CA72-4 alone and in combination for gastric cancer stratified by Hp infection status. Given that Hp infection may differentially affect the expression levels of these markers through various mechanisms, their diagnostic performance may vary across different Hp infection status. Thus, clarifying their diagnostic performance represents the first step toward optimizing gastric cancer screening strategies and customizing testing protocols based on Hp stratification.

Therefore, this study aims to evaluate the diagnostic value of serum HE4, TFF3, and CA72-4 alone and in combination for gastric cancer stratified by Hp infection status using a propensity score-matched case-control design, compare differences in diagnostic performance across Hp status, and investigate the incremental value of combined detection, thereby providing new theoretical evidence and clinical reference for early gastric cancer screening and individualized risk assessment.

2. Methods

2.1 Study Population

This was a single-center retrospective study conducted at Jinshazhou Hospital of Guangzhou University of Chinese Medicine. Data were retrieved from the hospital's de-identified electronic medical records and physical examination center database during routine clinical practice from January 2022 to December 2025.

Inclusion criteria of this study were as follows: (1) confirmed primary gastric adenocarcinoma by gastroscopic biopsy or surgical pathology; and (2) initial diagnosis with no prior anti-tumor therapy (radiotherapy, chemotherapy, targeted therapy, or immunotherapy).

Exclusion criteria included: (1) presence of other malignancies; (2) severe infection, autoimmune disease, or hepatic/renal dysfunction; (3) use of antibiotics, proton pump inhibitors, or bismuth agents within the past month (which may affect Hp testing accuracy); and (4) incomplete clinical data.

Individuals who underwent routine health checkups at the same hospital during the same period and had no prior history of any malignancy (especially no history of any gastrointestinal malignancy) were selected as the control group. Exclusion criteria for the control group included: (1) a history of gastric cancer or confirmed gastric cancer by examination; (2) a history of major gastrointestinal diseases (e.g., chronic atrophic gastritis, intestinal metaplasia); (3) severe infection, autoimmune disease, or hepatic/renal dysfunction; (4) use of antibiotics, proton pump inhibitors, or bismuth agents within the past month; and (5) incomplete clinical data.

2.2 Propensity Score Matching

To control for confounding bias, propensity score matching was performed to match cases and controls at a 1:1 ratio. Gastric cancer status was used as the dependent variable, with age, sex, body mass index (BMI), smoking history, alcohol consumption history, chronic kidney disease (CKD), chronic liver disease (CLD), and Hp infection status included as covariates in a logistic regression model to calculate propensity scores. Nearest neighbor matching was employed with a caliper width of 0.1, without replacement. After matching, covariate balance was assessed using standardized mean differences (SMD), with SMD <0.1 considered indicative of good between-group balance.

2.3 Data Collection and Definitions

Demographic and clinical data were retrospectively collected for all study subjects, including age, sex, body mass index (BMI), smoking history, alcohol consumption history, chronic kidney disease (CKD), chronic liver disease (CLD), and Hp infection status. For gastric cancer patients, additional tumor-related pathological information was collected, including tumor location, tumor–node–metastasis (TNM) stage (according to the American Joint Committee on Cancer [AJCC] 8th edition), histological differentiation grade, and Lauren classification for disease characterization. Peripheral venous blood samples were collected from all subjects prior to gastroscopy (for gastric cancer patients and control subjects who underwent the examination) or any anti-cancer therapy (for gastric cancer patients). For control subjects who did not undergo gastroscopy, samples were collected during their routine health checkup. Serum was separated and stored under standardized conditions. Enzyme-linked immunosorbent assay (ELISA) was employed for the measurements of HE4 (Human HE4 ELISA Kit, catalog No. EH4092, FineTest, Wuhan, China) and TFF3 (Human TFF3 ELISA Kit, catalog No. EH1378, FineTest, Wuhan, China) according to the manufacturer's instructions. CA72-4 was measured using electrochemiluminescence immunoassay (Roche Cobas e601, Roche Diagnostics, Mannheim, Germany). Hp infection status was determined by $^{13}C/^{14}C$ urea breath test (positive defined as delta over baseline [DOB] $\geq$4.0) or serum Hp antibody testing using ELISA for *Helicobacter pylori* immunoglobulin G (Hp-IgG), with positivity/negativity determined according to kit instructions.

2.4 Statistical Analysis

Statistical analyses were performed using R version 4.5.2 (R Foundation for Statistical Computing, Vienna, Austria). Continuous variables were first tested for normality using the Kolmogorov–Smirnov test; normally distributed variables were expressed as mean ± standard deviation, while non-normally distributed variables were expressed as median (interquartile range). Categorical variables were presented as counts (percentages). Between-group comparisons were performed using *t*-tests (for normally distributed continuous variables), Mann–Whitney *U* tests (for non-normally distributed continuous variables), or χ^2 tests/Fisher's exact tests (for categorical variables).

Study subjects were stratified according to Hp infection status into Hp-positive and Hp-negative strata. Within each stratum, marker levels were compared between the gastric cancer group and the control group.

Multivariate logistic regression analysis was performed to examine associations between serum markers and gastric cancer, with odds ratios (ORs) and 95% confidence intervals (CIs) calculated. Multicollinearity was assessed using the variance inflation factor (VIF). Receiver operating characteristic (ROC) curves were plotted for each marker, and area under the curve (AUC) values and their 95% CIs were calculated. Optimal cut-off values were determined using the Youden index at the optimal cut-off. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. DeLong's test was used to compare AUC differences between different markers or across strata. Combined diagnostic models were constructed using predicted probabilities from multivariate logistic regression, with models developed for the total population and each Hp stratum; combined AUC values were calculated and compared with those of single markers. Statistical significance was set at $p < 0.05$.

3. Results

3.1 Propensity Score Matching Outcomes

Initially, 357 gastric cancer patients and 3280 healthy controls were identified. After applying inclusion and exclusion criteria, 272 gastric cancer cases and 2000 controls were deemed eligible for analysis. Through propensity score matching, 268 gastric cancer patients were successfully matched with 268 healthy controls, yielding a final analytical sample of 536 subjects. Before matching, notable imbalances in smoking history, CKD, and Hp infection status were observed between the gastric cancer and control groups (SMD >0.1); after matching, all covariates achieved SMD <0.1, indicating good between-group balance (Table 1).

Table 1. Comparison of baseline characteristics between groups before and after matching.

Variable	Before matching				After matching			
	Gastric cancer group (*n* = 272)	Control group (*n* = 2000)	*p*	SMD	Gastric cancer group (*n* = 268)	Control group (*n* = 268)	*p*	SMD
Age (years), mean ± SD	57.86 ± 12.00	58.52 ± 10.96	0.360	0.057	57.86 ± 12.02	56.96 ± 11.11	0.369	0.078
Male, *n* (%)	172 (63.2)	1239 (62.0)	0.731	0.027	171 (63.8)	171 (63.8)	1	<0.001
BMI (kg/m^2), mean ± SD	23.05 ± 3.00	23.23 ± 2.93	0.347	0.060	23.07 ± 3.00	23.05 ± 2.93	0.933	0.007
Smoking history, *n* (%)	138 (50.7)	669 (33.5)	<0.001	0.356	134 (50.0)	135 (50.4)	1	0.007
Alcohol consumption history, n (%)	71 (26.1)	497 (24.9)	0.709	0.029	70 (26.1)	70 (26.1)	1	<0.001
CKD, *n* (%)	24 (8.8)	58 (2.9)	<0.001	0.254	20 (7.5)	24 (9.0)	0.637	0.054
CLD, *n* (%)	10 (3.7)	44 (2.2)	0.198	0.088	9 (3.4)	9 (3.4)	1	<0.001
Hp-positive, *n* (%)	173 (63.6)	796 (39.8)	<0.001	0.49	169 (63.1)	162 (60.4)	0.594	0.054

Abbreviations: BMI, body mass index; CKD, chronic kidney disease; CLD, chronic liver disease; Hp, *Helicobacter pylori*; SMD, standardized mean differences; SD, standard deviation.

3.2 Clinicopathological Characteristics of the Gastric Cancer Group

The clinicopathological characteristics of the 268 matched gastric cancer patients are presented in Table 2. Non-cardia tumors predominated (78.4%), the intestinal type was the most common Lauren classification (52.6%), and TNM stage III accounted for the largest proportion (42.5%). Regarding differentiation, poorly differentiated tumors accounted for the largest proportion (44.8%), followed by moderately differentiated tumors (41.0%), and well-differentiated tumors were the least common (14.2%).

Table 2. Clinicopathological characteristics of the gastric cancer group.

Variable	Number	Percentage (%)
Tumor location		
Cardia	58	21.6
Non-cardia	210	78.4
Lauren classification		
Intestinal type	141	52.6
Diffuse type	88	32.8
Mixed type	39	14.6
Differentiation grade		
Well-differentiated	38	14.2
Moderately differentiated	110	41.0
Poorly differentiated	120	44.8
TNM stage		
Stage I	45	16.8
Stage II	67	25.0
Stage III	114	42.5
Stage IV	42	15.7

Abbreviation: TNM, tumor–node–metastasis.

3.3 Comparison of Serum Marker Levels

In the total population, serum levels of HE4, TFF3, and CA72-4 were significantly higher in the gastric cancer group than in the control group (all $p < 0.001$). In the Hp-positive subgroup, all three markers were significantly elevated in the gastric cancer group compared with the control group (all $p < 0.001$). In the Hp-negative subgroup, HE4 and CA72-4 were significantly higher in the gastric cancer group ($p < 0.001$), and TFF3 also showed a statistically significant between-group difference ($p = 0.003$), though the magnitude of difference was smaller (Table 3).

3.4 Logistic Regression Analysis

Multicollinearity diagnostics showed no significant correlation among the three markers (all VIF <5). When all three markers were included simultaneously in a multivariate logistic regression model, HE4 (OR = 1.05, 95% CI: 1.03–1.06), TFF3 (OR = 1.19, 95% CI: 1.13–1.26), and CA72-4 (OR = 1.46, 95% CI: 1.34–1.62) remained independent risk factors for gastric cancer after adjustment (all $p < 0.001$) (Table 4).

3.5 Diagnostic Performance of Single and Combined Markers

In the total population, the AUC values for HE4, TFF3, and CA72-4 in diagnosing gastric cancer were 0.780 (0.740–0.820), 0.714 (0.671–0.758), and 0.782 (0.743–0.822), respectively. The combined model yielded an AUC of 0.899 (0.872–0.925), demonstrating significantly superior diagnostic performance compared with any individual marker (all $p < 0.001$).

After stratification by Hp status, marked differences in diagnostic performance emerged among the markers (Fig. 1). In the Hp-positive subgroup, TFF3 demonstrated the highest diagnostic efficacy (AUC = 0.825, 95% CI: 0.781–0.870), while HE4 and CA72-4 showed AUC values of 0.799 (0.749–0.849) and 0.777 (0.727–0.828), respectively; the combined model achieved an AUC of 0.941 (0.915–0.967). In the Hp-negative subgroup, CA72-4 showed the highest AUC (0.789, 95% CI: 0.725–0.852), followed by HE4 (0.749, 95% CI: 0.682–0.817); TFF3 exhibited markedly diminished performance (AUC = 0.620, 95% CI: 0.543–0.697). The combined model in Hp-negative individuals achieved an AUC of 0.866 (0.818–0.914).

DeLong's test revealed that the AUC of TFF3 was significantly higher in Hp-positive individuals than in Hp-negative individuals (0.825 vs. 0.620, $p < 0.001$). No significant difference was observed for CA72-4 (0.777 vs. 0.789, $p = 0.784$) and HE4 (0.799 vs. 0.749, $p = 0.245$) between the two strata. In the Hp-positive subgroup, the combined model (AUC = 0.941) exhibited significantly higher AUC than HE4 (0.799), TFF3 (0.825), and CA72-4 (0.777) (all $p < 0.001$ by DeLong's test). In the Hp-negative subgroup, the combined model (AUC = 0.866) also significantly outperformed each single marker (HE4: 0.749, $p < 0.001$; TFF3: 0.620, $p < 0.001$; CA72-4: 0.789, $p = 0.002$).

3.6 Diagnostic Performance Stratified by Hp Infection Status

Table 5 presents the optimal cut-off values, sensitivity, specificity, and AUC for each marker and the combined model in the overall population and Hp status-stratified subgroups. In the Hp-positive subgroup, TFF3 showed a substantially higher AUC (0.825, 95% CI: 0.781–0.870) than in the Hp-negative subgroup (0.620, 95% CI: 0.543–0.697), with an optimal cut-off of 12.95 ng/mL (sensitivity 68.0%, specificity 87.7%). In contrast, HE4 and CA72-4 demonstrated relatively stable diagnostic performance across Hp strata. The combined model achieved AUCs of 0.941 (0.915–0.967) in Hp-positive and 0.866 (0.818–0.914) in Hp-negative.

Table 3. Comparison of serum marker levels between gastric cancer and control groups in the total population and Hp status-stratified subgroups.

Stratum	Marker	Gastric cancer group	Control group	Z	*p*
Total population (*n* = 268)	HE4 (pmol/L)	66.6 (47.5–92.5)	42.6 (34.8–53.3)	11.22	<0.001
	TFF3 (ng/mL)	11.4 (7.2–17.3)	7.5 (5.5–9.9)	8.59	<0.001
	CA72-4 (U/mL)	8.5 (5.4–12.7)	4.7 (3.8–6.2)	11.31	<0.001
Hp-positive subgroup (control = 162; gastric cancer = 169)	HE4 (pmol/L)	75.6 (51.5–104.6)	43.0 (34.9–53.1)	9.40	<0.001
	TFF3 (ng/mL)	16.1 (11.0–22.3)	8.5 (6.7–11.5)	10.24	<0.001
	CA72-4 (U/mL)	8.7 (5.1–12.8)	4.8 (3.8–6.2)	8.72	<0.001
Hp-negative subgroup (control = 106; gastric cancer = 99)	HE4 (pmol/L)	60.5 (44.5–88.5)	41.6 (34.8–53.3)	6.16	<0.001
	TFF3 (ng/mL)	6.9 (5.3–9.1)	5.6 (4.5–7.4)	2.97	0.003
	CA72-4 (U/mL)	8.2 (5.4–12.4)	4.6 (3.7–6.2)	7.13	<0.001

Abbreviations: CA72-4, carbohydrate antigen 72-4; HE4, human epididymis protein 4; Hp, *Helicobacter pylori*; TFF3, trefoil factor 3.

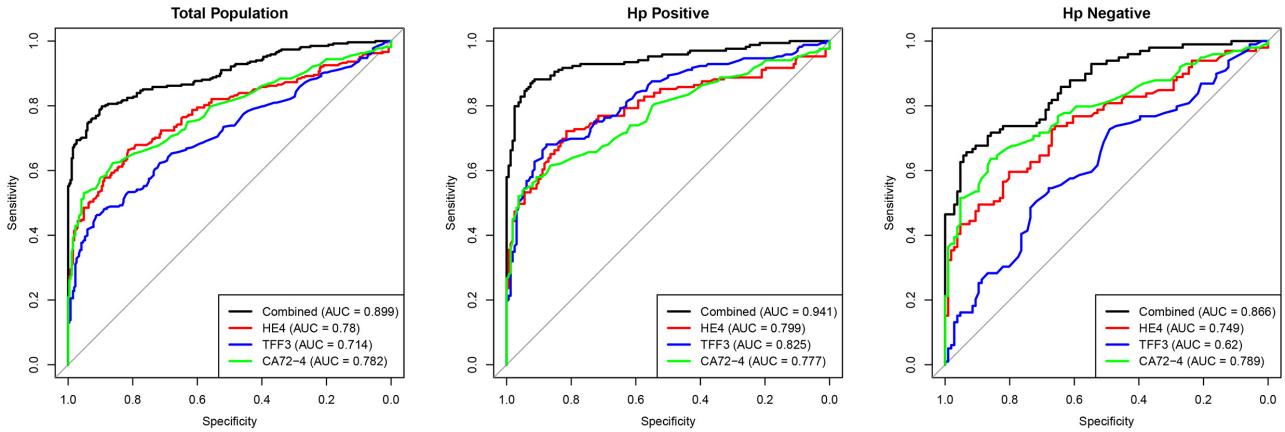

Fig. 1. Receiver operating characteristic (ROC) curves for serum markers in gastric cancer diagnosis. Abbreviation: AUC, area under the curve.

Table 4. Multivariate logistic regression analysis.

Predictor	Beta	SE	*p*	OR (95% CI)
HE4 (pmol/L)	0.045	0.006	<0.001	1.05 (1.03–1.06)
TFF3 (ng/mL)	0.173	0.028	<0.001	1.19 (1.13–1.26)
CA72-4 (U/mL)	0.381	0.049	<0.001	1.46 (1.34–1.62)

Notes: The multivariate logistic regression model included HE4, TFF3, and CA72-4 simultaneously. The adjustment refers to mutual adjustment among these three serum markers; no other covariates were entered into this model.

Abbreviations: CI, confidence interval; OR, odds ratio; SE, standard error.

4. Discussion

This study represents the first systematic evaluation of the diagnostic value of serum HE4, TFF3, and CA72-4, alone and in combination, for gastric cancer according to *Helicobacter pylori* (Hp) infection status. Our results demonstrate that all three markers are independent risk factors for gastric cancer, and combined detection achieves an AUC of 0.899, significantly outperforming any single marker. More importantly, we identified that Hp infection status significantly modifies the diagnostic performance of TFF3, while HE4 and CA72-4 remain relatively stable and unaffected by Hp status.

The AUC of TFF3 was significantly higher in the Hp-positive subgroup (0.825) than in the Hp-negative subgroup (0.620), with the difference reaching statistical significance ($p < 0.001$), indicating clear Hp-dependent diagnostic performance. Potential mechanisms underlying this phenomenon include the following: (1) Hp infection directly induces *TFF3* mRNA expression in gastric epithelial cells [13]; and (2) in intestinal metaplasia induced by Hp infection, goblet cells highly express TFF3, leading to further elevation of TFF3 in the context of Hp positivity. A systematic review by Zhang et al. [22] reported a pooled AUC of 0.80 for TFF3 in diagnosing gastric cancer, with significant associations with lymph node metastasis and TNM stage. Research by Zhao et al. [17] in patients with atrophic gastritis also found that the diagnostic value of TFF3 varied according to Hp status, corroborating our findings. Interestingly, a recent study revealed that chronic Hp infection promotes the selection of bacterial variants that effectively colonize metaplastic glands, a process that may further influence TFF3 expression in the gastric microenvironment [23], providing an evolutionary perspective on the close

Table 5. Diagnostic performance of individuals and combined markers in total population and Hp status-stratified subgroups.

Marker/Model	Subgroup	Cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC (95% CI)
HE4	Total	54.55	64.9	78.4	75.0	69.0	0.780 (0.740–0.820)
	Hp-positive	56.50	72.2	81.5	79.1	75.0	0.799 (0.749–0.849)
	Hp-negative	55.00	59.6	80.2	75.0	66.7	0.749 (0.682–0.817)
TFF3	Total	12.85	46.3	91.0	83.7	62.9	0.714 (0.671–0.758)
	Hp-positive	12.95	68.0	87.7	85.3	72.5	0.825 (0.781–0.870)
	Hp-negative	6.55	54.5	67.9	60.5	62.4	0.620 (0.543–0.697)
CA72-4	Total	6.85	62.3	86.2	81.9	69.5	0.782 (0.743–0.822)
	Hp-positive	8.15	53.8	95.1	91.9	66.8	0.777 (0.727–0.828)
	Hp-negative	6.95	63.6	85.8	79.7	73.2	0.789 (0.725–0.852)
Combined model	Total		79.9	89.6	88.7	82.7	0.899 (0.872–0.925)
	Hp-positive		87.0	92.6	90.0	85.4	0.941 (0.915–0.967)
	Hp-negative		65.7	93.4	93.9	76.5	0.866 (0.818–0.914)

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

link between Hp and TFF3. These lines of evidence suggest that TFF3 may have greater clinical utility for gastric cancer screening in Hp-positive populations than in Hp-negative populations. However, its superiority over other biomarkers or combined models in Hp-positive individuals cannot be concluded from the current data.

In our study, HE4 demonstrated an AUC of 0.799 (95% CI: 0.749–0.849) in the Hp-positive subgroup and 0.749 (95% CI: 0.682–0.817) in the Hp-negative subgroup, with DeLong's test showing no statistically significant difference between strata ($p = 0.245$). Tissue microarray studies by O'Neal et al. [24] demonstrated widespread HE4 overexpression in gastric cancer tissues (74% in intestinal-type, >90% in diffuse-type), with positive expression also observed in gastric metaplastic tissues, suggesting close associations between HE4 and both gastric cancer and precancerous lesions. Guo et al. [14] confirmed that HE4 promotes gastric cancer cell proliferation and migration through the Src/Akt/Erk signaling pathway, with high expression correlating with Lauren classification and TNM stage. A comprehensive review on the clinical utility of HE4 has recently summarized its emerging roles across multiple malignancies, including gastric cancer [25]. However, no studies to date have reported direct associations between HE4 expression or diagnostic performance and Hp infection status. The numerical difference observed in our study may reflect random variation due to the relatively smaller sample size in the Hp-negative subgroup. Thus, the diagnostic performance of HE4 is considered stable across different Hp status, a characteristic that renders it broadly applicable in combined detection strategies.

CA72-4 exhibited AUC values of 0.777 and 0.789 in Hp-positive and Hp-negative populations, respectively, with no statistically significant difference ($p = 0.784$), indicating that its diagnostic performance remains unaffected by Hp infection status. A systematic review by the Japanese Gastric Cancer Association. Shimada et al. [26] confirmed CA72-4 as one of the most commonly used serum markers for gastric cancer, with a positivity rate of approximately 30%, and demonstrated that this biomarker was significantly associated with tumor stage and patient survival. A meta-analysis by Wang et al. [19] demonstrated that CA72-4 combined with CEA and CA19-9 achieved an AUC of 0.87 for gastric cancer diagnosis. Despite the proven associations between Hp infection and elevated CA72-4 levels in healthy populations [20], our findings suggest that Hp infection does not significantly compromise the discriminatory ability of CA72-4 in gastric cancer diagnosis. This characteristic positions CA72-4 as a stable cornerstone marker in combined detection strategies.

In the total population and each of the Hp status-stratified subgroups, the AUC of combined detection significantly exceeded that of any single marker. Shen et al. [21] reported an AUC of 0.890 for a panel of five novel protein markers in diagnosing early gastric cancer, corroborating our findings and suggesting that multi-marker strategies represent an effective approach to overcoming the sensitivity limitations of single markers.

Based on these findings, we propose an Hp status-stratified testing strategy: in Hp-positive populations, TFF3 may serve as a preferred marker; in Hp-negative populations, the diagnostic value of TFF3 is limited, while both HE4 and CA72-4 can serve as stable components of combined testing; CA72-4 can serve as a foundational marker across all subgroups. The review by Sun [8] similarly emphasizes that Hp status-stratified testing strategies represent an important direction for enhancing diagnostic performance in the context of gastric cancer.

This study has several limitations. First, the retrospective design may introduce selection bias. Second, the single-center nature of the study population may limit the generalizability of our findings to other communities and populations. Third, external validation was not performed in this study, and associations between markers and gas-

tric cancer molecular subtypes were not explored. Fourth, the control group consisted exclusively of healthy individuals who sought physical examinations at the hospital, without including patients with gastritis, atrophic gastritis, or intestinal metaplasia (key precancerous lesions). This may introduce spectrum bias, potentially overestimating the diagnostic performance of the biomarkers. Future studies should include a precancerous lesion control group to better assess the biomarkers' value in early screening. Moreover, Hp infection status was determined using either the urea breath test or serum antibody testing. Antibody testing cannot distinguish between past and current infection, which may introduce misclassification bias. Although individuals with a recent history of eradication therapy were excluded from this study to better identify current infection, residual bias cannot be fully ruled out.

5. Conclusion

Serum HE4, TFF3, and CA72-4 are independent risk factors of gastric cancer, and their combined detection significantly enhances diagnostic efficacy. Hp infection status selectively influences the diagnostic value of these markers: TFF3 exhibits high diagnostic performance only in Hp-positive populations, while HE4 and CA72-4 demonstrate relatively stable diagnostic performance with no significant differences across Hp status. Therefore, Hp status-stratified testing strategies may help optimize marker selection, although further external validation is warranted before clinical implementation. This study provides new theoretical evidence supporting individualized auxiliary differential diagnosis of gastric cancer and suggests that Hp infection status should be considered when interpreting serum biomarker results in future clinical practice.

Key Points

• This study systematically evaluated the diagnostic value of serum HE4, TFF3, and CA72-4 for gastric cancer stratified by *Helicobacter pylori* (Hp) infection status.

• Hp infection status has a significant influence on the diagnostic performance of TFF3, which shows better accuracy (AUC = 0.825) in Hp-positive individuals than in Hp-negative individuals (AUC = 0.620), suggesting its potential utility in Hp-positive populations.

• HE4 and CA72-4 demonstrate stable diagnostic performance across different Hp status, with no significant differences between the Hp status-stratified subgroups.

• With superior diagnostic efficacy for gastric cancer regardless of Hp status (AUC = 0.899), the combined detection of HE4, CA72-4, and TFF3 provides a promising auxiliary tool for optimizing differential diagnosis of gastric cancer based on Hp status-stratified testing.

Availability of Data and Materials

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

Author Contributions

MYW designed the research study. MYW and MKX performed the research. MYW and MKX analyzed the data. MKX drafted the article. Both authors contributed to the important editorial changes in the manuscript. Both authors read and approved the final manuscript. Both authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

According to the Measures for Ethical Review of Biomedical Research Involving Humans and the regulations of our institutional ethics committee, the requirement to obtain informed consent was waived since this was a retrospective study that utilized anonymized clinical data. The study was approved by the Ethics Committee of Jinshazhou Hospital of Guangzhou University of Chinese Medicine (approval number: JSZ-IEC-SL-KT-20260224(1)), and all procedures were carried out in adherence to the principles of the Declaration of Helsinki.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: a Cancer Journal for Clinicians. 2021; 71: 209–249. https://doi.org/10.3322/caac.21660

[2] Li M, Cao S, Xu RH. Global trends and epidemiological shifts in gastrointestinal cancers: insights from the past four decades. Cancer Communications. 2025; 45: 774–788. https://doi.org/10.1002/cac2.70017

[3] Park JY, Georges D, Alberts CJ, Bray F, Clifford G, Baussano I. Global lifetime estimates of expected and preventable gastric cancers across 185 countries. Nature Medicine. 2025; 31: 3020–3027. https://doi.org/10.1038/s41591-025-03793-6

[4] Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F. Gastric cancer. Lancet. 2020; 396: 635–648. https://doi.org/10.1016/S0140-6736(20)31288-5

[5] Ajani JA, D'Amico TA, Almhanna K, Bentrem DJ, Chao J, Das P, et al. Gastric Cancer, Version 3.2016, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network. 2016; 14: 1286–1312. https://doi.org/10.6004/jnccn.2016.0137

[6] Orășeanu A, Brisc MC, Maghiar OA, Popa H, Brisc CM, Șolea SF, et al. Landscape of Innovative Methods for Early Diagnosis of Gastric Cancer: A Systematic Review. Diagnostics. 2023; 13: 3608. https://doi.org/10.3390/diagnostics13243608

[7] Feng F, Tian Y, Xu G, Liu Z, Liu S, Zheng G, et al. Diagnostic and prognostic value of CEA, CA19-9, AFP and CA125 for early gastric cancer. BMC Cancer. 2017; 17: 737. https://doi.org/10.1186/s12885-017-3738-y

[8] Sun HT. *Helicobacter pylori*-related serum indicators: Cutting-edge advances to enhance the efficacy of gastric cancer screening. World Journal of Gastrointestinal Oncology. 2025; 17: 100739. https://doi.org/10.4251/wjgo.v17.i3.100739

[9] Correa P. Human gastric carcinogenesis: a multistep and multifactorial process–First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. Cancer Research. 1992; 52: 6735–6740.

[10] Han Z, Liu J, Zhang W, Kong Q, Wan M, Lin M, et al. Cardia and non-cardia gastric cancer risk associated with Helicobacter pylori in East Asia and the West: A systematic review, meta-analysis, and estimation of population attributable fraction. Helicobacter. 2023; 28: e12950. https://doi.org/10.1111/hel.12950

[11] Riquelme A, Silva F, Reyes D, Latorre G. Chronic Atrophic Gastritis and Intestinal Metaplasia: A Latin American Perspective. The Korean Journal of Helicobacter and Upper Gastrointestinal Research. 2024; 24: 218–230. https://doi.org/10.7704/kjhugr.2024.0017

[12] Yang XT, Niu PQ, Li XF, Sun MM, Wei W, Chen YQ, et al. Differential cytokine expression in gastric tissues highlights helicobacter pylori's role in gastritis. Scientific Reports. 2024; 14: 7683. https://doi.org/10.1038/s41598-024-58407-x

[13] Matsuda K, Yamauchi K, Matsumoto T, Sano K, Yamaoka Y, Ota H. Quantitative analysis of the effect of Helicobacter pylori on the expressions of SOX2, CDX2, MUC2, MUC5AC, MUC6, TFF1, TFF2, and TFF3 mRNAs in human gastric carcinoma cells. Scandinavian Journal of Gastroenterology. 2008; 43: 25–33. https://doi.org/10.1080/00365520701579795

[14] Guo YD, Wang JH, Lu H, Li XN, Song WW, Zhang XD, et al. The human epididymis protein 4 acts as a prognostic factor and promotes progression of gastric cancer. Tumour Biology. 2015; 36: 2457–2464. https://doi.org/10.1007/s13277-014-2858-0

[15] Öksüz P, Diniz G, Akata T, Koca Y, Ayaz D. Correlation of Human Epididymis Protein 4 Expression with Clinicopathologic Parameters in Gastric Carcinomas. Hamidiye Medical Journal. 2025; 6: 59–66. https://doi.org/10.4274/hamidiyemedj.galenos.2025.15010

[16] Jing D, Jin J, Mei Z, Zhu Q, Lu Y, Wang X. Effects of *Helicobacter pylori* infection and interleukin 6 on the expression of ITIH4 in human gastric cancer cells. Translational Cancer Research. 2020; 9: 4656–4665. https://doi.org/10.21037/tcr-20-1766

[17] Zhao J, Tian W, Zhang X, Dong S, Shen Y, Gao X, et al. The diagnostic value of serum trefoil factor 3 and pepsinogen combination in chronic atrophic gastritis: a retrospective study based on a gastric cancer screening cohort in the community population. Biomarkers: Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals. 2024; 29: 384–392. https://doi.org/10.1080/1354750X.2024.2400927

[18] Latorre G, Pizarro M, Ford JS, Gándara V, Muñoz G, Araya JC, et al. Evaluation of trefoil factor 3 as a non-invasive biomarker of gastric intestinal metaplasia and gastric cancer in a high-risk population. Gastroenterologia Y Hepatologia. 2023; 46: 411–418. https://doi.org/10.1016/j.gastrohep.2022.04.004

[19] Wang H, Jin W, Wan C, Zhu C. Diagnostic value of combined detection of CA72-4, CA19-9, and carcinoembryonic antigen comparing to CA72-4 alone in gastric cancer: a systematic review and meta-analysis. Translational Cancer Research. 2022; 11: 848–856. https://doi.org/10.21037/tcr-22-537

[20] Hu PJ, Chen MY, Wu MS, Lin YC, Shih PH, Lai CH, et al. Clinical Evaluation of CA72-4 for Screening Gastric Cancer in A Healthy Population: A Multicenter Retrospective Study. Cancers. 2019; 11: 733. https://doi.org/10.3390/cancers11050733

[21] Shen W, Zhou H, Li L, Liu W, Lou Q, Tong CY, et al. Promising protein biomarkers for early gastric cancer: clinical performance of combined detection. Clinical Chemistry and Laboratory Medicine. 2025; 63: 1864–1875. https://doi.org/10.1515/cclm-2024-1510

[22] Zhang CX, Wu CT, Xiao L, Tang SH. The diagnostic and clinicopathological value of trefoil factor 3 in patients with gastric cancer: a systematic review and meta-analysis. Biomarkers: Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals. 2021; 26: 95–102. https://doi.org/10.1080/1354750X.2020.1871411

[23] O'Brien VP, Jackson LK, Frick JP, Rodriguez Martinez AE, Jones DS, Johnston CD, et al. Helicobacter pylori Chronic Infection Selects for Effective Colonizers of Metaplastic Glands. mbio. 2023; 14: e0311622. https://doi.org/10.1128/mbio.03116-22

[24] O'Neal RL, Nam KT, LaFleur BJ, Barlow B, Nozaki K, Lee HJ, et al. Human epididymis protein 4 is up-regulated in gastric and pancreatic adenocarcinomas. Human Pathology. 2013; 44: 734–742. https://doi.org/10.1016/j.humpath.2012.07.017

[25] Qiu M, Zou S, Yu Z, Nian X. Clinical utility of human epididymis protein 4. Critical Reviews in Clinical Laboratory Sciences. 2025; 62: 510–534. https://doi.org/10.1080/10408363.2025.2512467

[26] Shimada H, Noie T, Ohashi M, Oba K, Takahashi Y. Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the Task Force of the Japanese Gastric Cancer Association. Gastric Cancer. 2014; 17: 26–33. https://doi.org/10.1007/s10120-013-0259-5