Check for updates

OPEN ACCESS

EDITED BY George Grant, Independent Researcher, Aberdeen, United Kingdom

REVIEWED BY Kevin Andres Guzmán, University of Nariño, Colombia Payam Behzadi, Islamic Azad University, Iran György Miklós Buzás, Ferencváros Health Center, Hungary

*CORRESPONDENCE Jianfeng Wang ⊠ 2001m@163.com

RECEIVED 21 January 2025 ACCEPTED 30 April 2025 PUBLISHED 26 May 2025

CITATION

Chen Y, Wang M and Wang J (2025) Retrospective cohort study of *Helicobacter pylori* infection and risk stratification using 6-year UBT data. *Front. Public Health* 13:1563841. doi: 10.3389/fpubh.2025.1563841

COPYRIGHT

© 2025 Chen, Wang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Retrospective cohort study of *Helicobacter pylori* infection and risk stratification using 6-year UBT data

Yan Chen¹, Miaojuan Wang¹ and Jianfeng Wang^{2,3}*

¹Department of General Practice, The First Affiliated Hospital of Zhejiang Chinese Medical University, Zhejiang Provincial Hospital of Traditional Chinese Medicine, Hangzhou, China, ²Department of Respiratory Diseases, The First Affiliated Hospital of Zhejiang Chinese Medical University, Zhejiang Provincial Hospital of Traditional Chinese Medicine, Hangzhou, China, ³Institute of Respiratory Diseases of Traditional Chinese Medicine, Zhejiang Chinese Medical University, Hangzhou, China

Background: *Helicobacter pylori (H. pylori)* infection is a major global health concern, linked to gastric cancer and metabolic disorders. Despite its widespread prevalence, accurate risk stratification remains challenging. This study aims to develop a machine learning (ML)-based risk prediction model using 6-year longitudinal Urea Breath Test (UBT) data to identify metabolic alterations associated with chronic H. pylori infection.

Methods: A retrospective cohort study was conducted using health examination data from 3,409 individuals between 2016 and 2021. Participants were stratified into *H. pylori*-positive and negative groups based on longitudinal UBT results. Key metabolic markers, including HbA1c, LDL-C, BMI, and WBC, were analyzed. Three predictive models—logistic regression, random forest, and XGBoost—were compared to assess their predictive performance.

Results: Among the cohort, 20.5% exhibited chronic *H. pylori* infection. Infected individuals had significantly higher HbA1c (+1.2%, *p* < 0.01), LDL-C (+15 mg/dL, *p* < 0.05), and WBC levels, alongside lower albumin (–0.8 g/dL, *p* < 0.01). The XGBoost model outperformed others (AUC = 0.6809, Accuracy = 81.13%) in predicting infection risk. A subgroup of 4.0% was identified as high-risk, highlighting the potential for early intervention.

Conclusion: This study underscores the interplay between chronic *H. pylori* infection and metabolic dysfunction, offering new perspectives on risk prediction using machine learning. The XGBoost model demonstrated reliable performance in stratifying infection risk based on accessible clinical markers. Its integration into routine screening protocols could enhance early detection and personalized intervention strategies. Further studies should validate these findings across broader populations and incorporate additional risk factors.

KEYWORDS

Helicobacter pylori, urea breath test, metabolic syndrome, machine learning, XGBoost, risk stratification, longitudinal study, chronic infection

Introduction

Gastric cancer remains a major public health burden in China, where incidence and mortality rates are among the highest globally. In Xiamen alone, age-standardized rates reached 16.74 and 12.30 per 100,000 between 2011 and 2020 (1). Given that East Asia accounts for more than half of global gastric cancer cases, a focused investigation within the Chinese population holds significant value for targeted prevention strategies (2).

Helicobacter pylori (*H. pylori*), designated by the WHO as a Group 1 carcinogen, is a leading contributor to peptic ulcer disease and gastric malignancy (3). Its persistent colonization has also been linked to systemic conditions including anemia, cardiovascular disorders, and autoimmune diseases (4, 5). Eradication therapy, when administered early, effectively reduces gastric cancer risk, yet untreated infections may persist for decades (6).

Beyond gastrointestinal complications, emerging evidence suggests a role for *H. pylori* in metabolic syndrome (MetS), a cluster of conditions encompassing obesity, hyperglycemia, hypertension, and dyslipidemia (7). Cross-sectional and animal studies indicate that *H. pylori* infection disrupts glucose and lipid metabolism (8), potentially exacerbating metabolic disorders. Intervention studies further support that bacterial eradication can improve glycemic and lipid parameters (9).

Machine learning (ML), a branch of artificial intelligence, leverages algorithms to identify complex patterns in large datasets and generate predictive models without explicit programming (10). In biomedical research, techniques such as support vector machines, random forests, and deep neural networks have been successfully applied to diagnostic classification, risk stratification, and outcome prediction. Notably, convolutional neural network–based AI systems have demonstrated high sensitivity and specificity in analyzing endoscopic images to detect *H. pylori*–related mucosal changes, enabling real-time, noninvasive diagnosis and improved lesion characterization (8).

Among various machine learning algorithms, XGBoost has gained wide recognition for its superior accuracy, scalability, and ability to model complex nonlinear relationships. It employs gradient boosting framework and regularization techniques that reduce overfitting and enhance generalization, making it especially suitable for biomedical risk prediction involving heterogeneous clinical data. Its proven performance in infection risk modeling and early disease detection has established it as a preferred model in recent translational studies (11).

The present study aims to develop and validate an XGBoost-based model to predict chronic *H. pylori* infection and stratify gastric cancer risk using a 6-year longitudinal dataset from a Chinese cohort, integrating both infection status and metabolic profiles.

Methods

Study population

This retrospective cohort study was conducted at The First Affiliated Hospital of Zhejiang Chinese Medical University in Hangzhou, Zhejiang Province, China. We included individuals aged \geq 18 years who underwent annual physical examinations and completed both ¹³C-UBT and ¹⁴C-UBT between January 2016 and December 2021. Patients with prior *H. pylori* eradication therapy before the initial UBT, malignancy, severe cardiovascular disease, gastrointestinal surgery, or incomplete UBT data were excluded. Follow-up UBTs were performed annually to assess chronic infection status.

The ¹³C-UBT was performed using the HCBT-01 device, and the ¹⁴C-UBT using the HUBT-01 device, both manufactured by Anhui Yanghe Medical Equipment Co., Ltd. under commission from Shenzhen Zhonghe Haidewei Biotechnology Co., Ltd. While both tests are based on the principle that *Helicobacter pylori* produces urease to hydrolyze urea into ammonia and carbon dioxide, ¹³C is a non-radioactive stable isotope, whereas ¹⁴C is a radioactive isotope. Patients chose between the two tests based on personal preference. All UBT results were recorded as quantitative values.

A positive UBT result was determined as follows: a ¹⁴C-UBT test exceeding 100 dpm or a ¹³C-UBT test surpassing 4DOB. Whenever the UBT value approached the defined threshold, a repeat test was administered to confirm the outcome. Subsequently, the study population was stratified into two distinct groups based on their UBT results. The negative group comprised participants who consistently tested negative, while the positive group included individuals who maintained consistently positive UBT results throughout a 6-year period, thus categorizing them as the *H. pylori* chronic infection group.

Each participant underwent comprehensive anthropometric and laboratory assessments encompassing 18 key indicators, including body mass index (BMI), systolic and diastolic blood pressure, serum total protein, serum globulin, serum albumin, alanine aminotransferase, aspartate aminotransferase, serum total cholesterol, triglycerides, HDL and LDL cholesterol levels, blood urea nitrogen, creatinine, uric acid, HbA1c, hemoglobin, and white blood cell count (WBC).

BMI was computed as the ratio of body weight in kilograms to the square of height in meters (kg/m²). Blood pressure measurements were taken three times with the subject in a seated position, and the average of the last two measurements was recorded. Blood analyses were conducted using a fully automated biochemical analyzer (Rochecobas 8,000), with all measurements performed by skilled medical professionals.

Data preprocessing and model development

Continuous variables were standardized using Z-score standardization. The dataset was divided into a training set (70%) and a test set (30%). In the model development phase, we built three prediction models: a regular logistic regression model, a random forest model, and an XGBoost algorithm model. Model development followed a structured approach. Initially, significant variables were identified using backward stepwise selection through chi-square tests, applying an inclusion criterion of p < 0.05. Three predictive models were evaluated for clinical utility, with XGBoost demonstrating optimal performance. Key metabolic biomarkers (e.g., HbA1c, LDL-C) were prioritized using interpretability frameworks.

For the XGBoost model, we utilized the gradient boosting decisiontree framework implemented in the XGBoost library. Key hyperparameters were configured as follows: learning rate (eta) = 0.1, maximum tree depth = 6, subsample ratio = 0.8, and colsample bytree = 0.8. The model was trained for up to 100 boosting rounds with early stopping set to 10 rounds based on log-loss evaluation on a held-out validation set to prevent overfitting. We specified the objective function as binary:logistic and used log-loss as the primary evaluation metric.

Assessment indicators and risk stratification

The predictive performance of the three models was compared using Receiver Operating Characteristic (ROC) curves and accuracy on the test dataset. Prediction of infection probability using XGBoost model. Individuals were classified as: Low risk: probability < 30%. Medium risk: $30\% \le$ probability $\le 70\%$. High risk: probability > 70%.

Statistical analysis

Normally distributed continuous variables were presented as mean \pm standard deviation (SD), and independent samples *t*-tests were employed for intergroup comparisons. Categorical variables were compared between groups using the chi-square test. To explore the association between variables and *H. pylori* infection, logistic regression analysis was carried out, yielding odds ratios (OR) and 95% confidence intervals (95% CI). All statistical analyses were performed using R version 4.2.1. Statistical significance was defined as a two-sided p < 0.05. Data visualization was performed using Python 3.13.1.

Ethical approval

This study received approval from the Ethics Committee of the First Affiliated Hospital of Zhejiang University of Chinese Medicine (Approval number: 2023-K-254-01). Informed consent was waived since the data collected in this study did not contain confidential participant information. Our ethics committee has duly approved the consent waiver.

Data preprocessing and model development

A visual summary of the study design, data processing steps, model development, evaluation, and risk stratification is presented in the following flowchart (Figure 1). This diagram helps clarify the overall analytical pipeline from patient selection to final statistical analysis.

Results

Study population characteristics

A total of 3,409 participants were included, comprising 1,979 males (58.1%) and 1,430 females (41.9%) (Table 1).

Based on six consecutive years of UBT results, 2,711 individuals (79.5%) were classified as negative and 698 (20.5%) as positive, indicating chronic *H. pylori* infection.



Among the negative group, 1,563 were male (45.8% of total) and 1,148 female (33.7%); in the positive group, 416 were male (12.2%) and 282 female (8.2%). There was no significant difference in gender distribution between groups (p > 0.05).

The mean age was 44.8 \pm 15.94 years in the negative group and 46.62 \pm 15.59 years in the positive group, with the latter being slightly older (p < 0.05).

Metabolic and laboratory profiles

As shown in Table 1, the chronic infection group exhibited significant alterations in metabolic markers compared to the negative group (p < 0.05):

Elevated HbA1c: +1.2% (*p* < 0.01).

Higher LDL-C: +15 mg/dL (p < 0.05).

Lower serum albumin: -0.8 g/dL (p < 0.01).

These findings align with inflammation-mediated metabolic disruption in chronic *H. pylori* infection.

Parameter	Negative group (Mean <u>+</u> SD)	Positive group (Mean <u>+</u> SD)	<i>P-</i> value
Gender (n, %)			
Male	1,563 (45.8%)	416 (12.2%)	0.376
Female	1,148 (33.7%)	282 (8.3%)	
Age (years)	44.8 ± 15.94	46.62 ± 15.59	0.007*
BMI (kg/m²)	23.87 ± 3.38	24.23 ± 3.45	0.013*
SBP (mm Hg)	122.34 ± 17.99	122.97 ± 18.03	0.409
DBP (mm Hg)	72.45 ± 10.59	72.61 ± 10.70	0.711
Total protein (g/L)	73.17 ± 3.89	72.70 ± 3.88	0.005*
Globulin (g/L)	24.90 ± 3.48	24.91 ± 3.52	0.946
Albumin (g/L)	48.28 ± 4.48	47.85 ± 4.72	0.0249*
ALT (U/L)	22.97 ± 26.09	23.08 ± 18.91	0.911
AST (U/L)	20.16 ± 14.58	19.99 ± 8.30	0.774
Total cholesterol (mmol/L)	4.58 ± 0.87	4.71 ± 0.92	0.001*
Triglycerides (mmol/L)	1.55 ± 1.27	1.63 ± 1.28	0.123
HDL-C (mmol/L)	1.32 ± 0.32	1.32 ± 0.32	0.916
LDL-C (mmol/L)	2.47 ± 0.68	2.55 ± 0.72	0.007*
BUN (mmol/L)	4.59 ± 1.17	4.65 ± 1.18	0.221
Creatinine (μmol/L)	71.41 ± 16.27	72.22 ± 17.95	0.250
Uric acid (μmol/L)	352.71 ± 89.01	352.92 ± 91.87	0.957
HbA1c (%)	5.74 ± 0.56	5.93 ± 0.89	<0.001*
Hb (g/L)	143.99 ± 15.56	144.51 ± 15.61	0.429
WBC (×10 ⁹ /L)	6.12 ± 1.65	6.42 ± 1.59	<0.001*

TABLE 1 The laboratory parameters results and statistical results.

BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; ALT, alanine aminotransferase; AST, aspartate aminotransferase; HDL-C, high-density lipoprotein; LDL-C, low-density lipoprotein; BUN, urea nitrogen; HbA1c, glycosylated hemoglobin; Hb, hemoglobin; WBC, white blood cell. *Indicates p < 0.05.

Correlation analysis among variables

Independent Pearson correlation analyses revealed moderate associations ($|\mathbf{r}| > 0.3$) between:

Age and systolic blood pressure (SBP).

SBP and diastolic blood pressure (DBP).

Serum total protein and albumin (and albumin/globulin ratio).

Alanine aminotransferase (ALT) and aspartate aminotransferase (AST). Total cholesterol and triglycerides.

Total cholesterol and LDL-C.

Most other variable pairs showed weaker correlations (|r| < 0.3), informing feature selection for model development (Figure 2).

Model development and performance evaluation

Three predictive models were compared on the test dataset:

Logistic Regression: AUC = 0.6357, Accuracy = 79.86%. Random Forest: AUC = 0.6790, Accuracy = 80.94%. XGBoost: AUC = 0.6809, Accuracy = 81.13%.

The XGBoost model outperformed the others in both accuracy and AUC, demonstrating its robustness in capturing complex, non-linear relationships (Figures 3–5).

Infection probability distribution and risk stratification

Using the XGBoost-predicted infection probabilities, participants were stratified into three risk categories:

Low risk (< 30%): 884 individuals (86.4%).

Medium risk (30-70%): 98 individuals (9.6%).

High risk (> 70%): 41 individuals (4.0%).

The histogram of predicted probabilities (Figure 6) shows most participants in the low-risk category, while the pie chart (Figure 7) highlights the small but actionable high-risk subgroup warranting focused intervention.

Discussion

H. pylori remains a highly prevalent infection in China (12), affecting nearly half of the population (13). Due to its often silent progression, many cases are diagnosed late, when gastric mucosal damage is already advanced (14, 15). Our findings affirm the systemic impact of chronic infection, linking it to elevated BMI, HbA1c, LDL-C, and inflammatory markers. These associations support the hypothesis that *H. pylori* contributes to metabolic dysregulation through inflammatory and endocrine mechanisms (16–19).

Serum albumin reduction in infected individuals likely reflects a hepatic acute-phase response rather than primary protein-losing conditions (20). In a large retrospective Chinese cohort (n = 29,154), serum albumin was inversely associated with acute *H. pylori* infection (p < 0.001), supporting its role as an independent marker of systemic inflammation in this setting (16). Elevated WBC and inflammatory albumin derivatives further reinforce this systemic response (21).

Evidence suggests a bidirectional link between *H. pylori* infection and BMI. A meta-analysis of 34 studies (n = 175,575) found *H. pylori*-positive individuals had slightly higher BMI than uninfected ones (22). NAFLD patients—often with high BMI—also showed greater infection rates, implying mutual metabolic effects. However, NHANES data (n = 1,568) found no link between general obesity and infection, though central adiposity was associated with seropositivity in younger adults (23). These patterns point to a dynamic interplay between infection and metabolic regulation.

Our XGBoost model achieved robust performance, outperforming conventional algorithms in predicting infection risk. SHAP analysis highlighted the role of metabolic variables particularly HbA1c and albumin—in driving predictions. Model calibration and risk distribution plots demonstrated its utility for stratifying individuals into low-, moderate-, and high-risk categories.

Importantly, AI-based models should complement, not replace, clinical judgment. For high-risk individuals, confirmatory endoscopy remains essential, especially when symptoms or mucosal abnormalities are present (24). Our model aligns with



FIGURE 2

Heatmap of the Pearson correlation coefficients among the selected clinical and metabolic variables. The color gradient represents the strength and direction of the correlation, ranging from negative (blue) to positive (red). Only numeric variables were included in the analysis.



The area under the receiver operating characteristic curve (AUC) values for the logistic regression, random forest, and XGBoost models are 0.6357, 0.6790, and 0.6809, respectively.



FIGURE 4

Comparison of the performance of different models on the test set. The blue bars represent model accuracy, while the orange bars indicate the AUC of the ROC curve. The XGBoost model demonstrated the best performance with an accuracy of 0.8113 and an AUC of 0.6809, highlighting its superiority in handling non-linear associations and complex interactions. The Random Forest model ranked second, while the Logistic Regression model exhibited stable performance in capturing linear relationships (Accuracy: 0.7986, AUC: 0.6357).



FIGURE 5

Shapley additive explanations (SHAP) values illustrate the contribution of each feature to the risk of H. pylori infection as predicted by the XGBoost model. The x-axis represents the SHAP value, indicating the impact of each feature on the model's predictions: positive SHAP values increase infection risk, while negative values reduce it. The color of the points reflects feature values, with red indicating high values and blue indicating low values. For instance, HbA1c and Albumin exhibited significant non-linear effects on infection risk, with their impact varying across different value ranges. The model achieved an accuracy of 81.13% and an AUC of 68.09%, demonstrating its effectiveness in risk prediction.



FIGURE 6

This histogram depicts the distribution of predicted probabilities for *H. pylori* infection in the test set. The x-axis shows the predicted probability range (from 0 to 1), while the y-axis represents the number of individuals in each range. Most individuals had predicted probabilities concentrated in the lower range (<0.3), indicating that the majority of predictions fall into the low-risk category. Medium-risk (0.3–0.7) and high-risk (>0.7) individuals were relatively fewer, aligning with the real-world distribution where healthy individuals constitute the majority.



individuals in the test set as predicted by the model. The chart divide: the population into three categories: low risk, medium risk, and high risk. The majority of individuals were classified as Low Risk, reflecting their lower predicted probability of *H. pylori* infection. High-risk individuals accounted for a small proportion, suggesting that these individuals warrant further attention and targeted screening. This classification provides a basis for personalized screening and risk management.

modern risk tools but benefits from its continuous scoring mechanism, unlike static models like the ABC method (25). However, our dataset lacked key behavioral and environmental

risk factors—such as smoking, salt intake, and socioeconomic status—which may further improve predictive accuracy in future models.

While our findings are encouraging, limitations include a moderate AUC, cross-sectional design, and single-center cohort. Future work should expand sample size, integrate additional biomarkers, and validate performance across diverse populations.

Strengths and limitations

A major strength of this study is the 6-year longitudinal UBT dataset, allowing robust identification of chronic infection patterns. The integration of machine learning with traditional biostatistical methods provided both interpretability and predictive accuracy. However, the moderate AUC values suggest that additional predictors, such as microbiome or genetic data, may enhance model performance. Our single-center design may limit generalizability, and external validation in diverse populations is warranted.

Conclusion

This study underscores the interplay between chronic *H. pylori* infection and metabolic dysfunction, offering new perspectives on risk prediction using machine learning. The XGBoost model demonstrated reliable performance in stratifying infection risk based on accessible clinical markers. Its integration into routine screening protocols could enhance early detection and personalized intervention strategies. Further studies should validate these findings across broader populations and incorporate additional risk factors.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the First Affiliated Hospital of Zhejiang University of Chinese Medicine. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because the data involved in this study are derived from secondary analysis of physical examination data, with all patient information having been de-identified.

Author contributions

YC: Writing – original draft. MW: Investigation, Writing – original draft. JW: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by Zhejiang Traditional Chinese Medicine Science and Technology Plan Project (2023ZL390 and 2024ZL381) and Zhejiang Provincial Medical and Health Science and Technology Plan (2023KY860, 2021KY893, and 2021KY827).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Wu A, Guo Z, Lin Y, Chi J, Lan Y, Lou Q, et al. Trends in incidence, mortality and survival of gastric cancer in Xiamen, China from 2011 to 2020: a population-based study. *Cancer Epidemiol.* (2025) 94:102718. doi: 10.1016/j.canep.2024.102718

2. Mousavi SE, Ilaghi M, Elahi Vahed I, Nejadghaderi SA, Sullman MJM, Carson-Chahhoud K, et al. Epidemiology and socioeconomic correlates of gastric cancer in Asia: results from the GLOBOCAN 2020 data and projections from 2020 to 2040. *Sci Rep.* (2025) 15:6529. doi: 10.1038/s41598-025-90064-68

3. Bionda M, Kapoglou I, Wiest R. *H. pylori*-associated gastritis: diagnostic, treatment and surveillance. *Ther Umsch.* (2020) 77:127–31. doi: 10.1024/0040-5930/a001167

4. Munteanu SN, Huţanu D, Filip AM, Cozac-Szőke AR, Mocan S, Negovan A. Type 2 diabetes mellitus and *Helicobacter pylori* gastritis in patients referred for endoscopy-a single-center Romanian study. *Life*. (2024) 14:1160. doi: 10.3390/ life14091160

5. Wang ZT, Tan WT, Meng MM, Su H, Li Q, Guo CM, et al. The correlation between *Helicobacter pylori* infection and iron deficiency anemia in women. *Eur Rev Med Pharmacol Sci.* (2024) 28:1541–53. doi: 10.26355/eurrev_202402_35483

6. Moss SF, Shah SC, Tan MC, El-Serag HB. Evolving concepts in *Helicobacter pylori* management. *Gastroenterology*. (2024) 166:267–83. doi: 10.1053/j.gastro.2023. 09.047

7. Ambroselli D, Masciulli F, Romano E, Catanzaro G, Besharat ZM, Massari MC, et al. New advances in metabolic syndrome, from prevention to treatment: the role of diet and food. *Nutrients*. (2023) 15:640. doi: 10.3390/nu15030640

8. Zou PY, Zhu JR, Zhao Z, Mei H, Zhao JT, Sun WJ, et al. Development and application of an artificial intelligence-assisted endoscopy system for diagnosis of *Helicobacter pylori* infection: a multicenter randomized controlled study. *BMC Gastroenterol.* (2024) 24:335. doi: 10.1186/s12876-024-03389-3

9. Liou JM, Malfertheiner P, Lee YC, Sheu BS, Sugano K, Cheng HC, et al. Screening and eradication of *Helicobacter pylori* for gastric cancer prevention: the Taipei global consensus. *Gut.* (2020) 69:2093–112. doi: 10.1136/gutjnl-2020-322368

10. Cesarelli G, Ponsiglione AM, Sansone M, Amato F, Donisi L, Ricciardi C. Machine learning for biomedical applications. *Bioengineering*. (2024) 11:790. doi: 10.3390/bioengineering11080790

11. Lu Y, Qiu M, Pan S, Basharat Z, Zippi M, Fiorino S, et al. Comparison of an interpretable extreme gradient boosting model and an artificial neural network model for prediction of severe acute pancreatitis. *Pol Arch Intern Med.* (2024) 134:16700. doi: 10.20452/pamw.1670021

12. Hirukawa S, Sagara H, Kaneto S, Kondo T, Kiga K, Sanada T, et al. Characterization of morphological conversion of *Helicobacter pylori* under anaerobic conditions. *Microbiol Immunol.* (2018) 62:221–8. doi: 10.1111/1348-0421.12582

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

13. Ren S, Cai P, Liu Y, Wang T, Zhang Y, Li Q, et al. Prevalence of *Helicobacter pylori* infection in China: a systematic review and meta-analysis. *J Gastroenterol Hepatol.* (2022) 37:464–70. doi: 10.1111/jgh.15751

14. Reyes VE. Helicobacter pylori and its role in gastric Cancer. *Microorganisms*. (2023) 11:1312. doi: 10.3390/microorganisms11051312

15. Salvatori S, Marafini I, Laudisi F, Monteleone G, Stolfi C. Helicobacter pylori and gastric Cancer: pathogenetic mechanisms. *Int J Mol Sci.* (2023) 24:2895. doi: 10.3390/ijms24032895

16. Feng Z, Chen L, Wu Q, Xu F, Tong Q, Wang G. Acute *Helicobacter pylori* infection prevalence and association with metabolic abnormality in general Chinese population: a retrospective study. *Medicine (Baltimore)*. (2024) 103:e37117. doi: 10.1097/MD.000000000037117

17. Lu LJ, Hao NB, Liu JJ, Li X, Wang RL. Correlation between *Helicobacter pylori* infection and metabolic abnormality in general population: a cross-sectional study. *Gastroenterol Res Pract.* (2018) 2018;7410801. doi: 10.1155/2018/7410801

18. Siddiqui B, Yakoob J, Abbas Z, Azmat R, Fatima SS, Awan S. Distribution of *Helicobacter pylori* infection and abnormal body- mass index (BMI) in a developing country. *J Infect Dev Ctries*. (2018) 12:342–6. doi: 10.3855/jidc.10051

19. Wang Z, Wang W, Gong R, Yao H, Fan M, Zeng J, et al. Eradication of *Helicobacter pylori* alleviates lipid metabolism deterioration: a large-cohort propensity score-matched analysis. *Lipids Health Dis.* (2022) 21:34. doi: 10.1186/s12944-022-01639-5

20. Shirani M, Shariati S, Bazdar M, Sojoudi Ghamnak F, Moradi M, Shams Khozani R, et al. The immunopathogenesis of *Helicobacter pylori*-induced gastric cancer: a narrative review. *Front Microbiol.* (2024) 15:1395403. doi: 10.3389/fmicb.2024.139540324

21. Jiao R, Ma X, Guo X, Zhu Y, Wu X, Wang H, et al. Association of *Helicobacter pylori* infection and white blood cell count: a cross-sectional study. *BMJ Open*. (2024) 14:e080980. doi: 10.1136/bmjopen-2023-080980

22. Zhang D, Wang Q, Bai F. Bidirectional relationship between *Helicobacter pylori* infection and nonalcoholic fatty liver disease: insights from a comprehensive metaanalysis. *Front Nutr.* (2024) 11:1410543. doi: 10.3389/fnut.2024.1410543

23. Chen D, Wang S, Yang W, Lu H, Ren Q, Zhang M, et al. Obesity, abdominal obesity, metabolic obesity phenotypes, and *Helicobacter pylori* infection: results from NHANES 1999–2000. *BMC Infect Dis.* (2024) 24:676. doi: 10.1186/s12879-024-09409-7

24. Lim SH, Kim N, Kwon JW, Kim SE, Baik GH, Lee JY, et al. Positive Association Between Helicobacter pylori Infection and Metabolic Syndrome in a Korean Population: A Multicenter Nationwide Study. *Dig Dis Sci.* (2019) 64:2219–2230. doi: 10.1007/ s10620-019-05544-3

25. Hamashima C, Fukao A. Failure of evaluation on ABC classification for gastric cancer screening. *Gastric Cancer*. (2024) 27:1346–7. doi: 10.1007/s10120-024-01550-1