REVIEW
1604 Advances in postoperative adjuvant therapy for primary liver cancer
Zeng ZM, Mo N, Zeng J, Ma FC, Jiang YF, Huang HS, Liao XW, Zhu GZ, Ma J, Peng T

1622 Immunotherapy for nonalcoholic fatty liver disease-related hepatocellular carcinoma: Lights and shadows
Costante F, Airola C, Santopao F, Gasbarrini A, Pompili M, Ponziani FR

1637 Emerging role of caldesmon in cancer: A potential biomarker for colorectal cancer and other cancers

MINIREVIEWS
1654 Liquid biopsy to detect resistance mutations against anti-epidermal growth factor receptor therapy in metastatic colorectal cancer
Valenzuela G, Burotto M, Marcelain K, González-Montero J

1665 Implication of gut microbiome in immunotherapy for colorectal cancer

ORIGINAL ARTICLE

Basic Study
1675 Potential of six-transmembrane epithelial antigen of the prostate 4 as a prognostic marker for colorectal cancer
Fang ZX, Li CL, Chen WJ, Wu HT, Liu J

Case Control Study
1689 Inverse relations between Helicobacter pylori infection and risk of esophageal precancerous lesions in drinkers and peanut consumption

Retrospective Cohort Study
1699 Prognostic impact of tumor deposits on overall survival in colorectal cancer: Based on Surveillance, Epidemiology, and End Results database
Wu WX, Zhang DK, Chen SX, Hou ZY, Sun BL, Yao L, Jie JZ

1711 Consolidation chemotherapy with capecitabine after neoadjuvant chemoradiotherapy in high-risk patients with locally advanced rectal cancer: Propensity score study
Monthly Volume 14 Number 9 September 15, 2022

**Retrospective Study**

1727  Efficacy and safety of computed tomography-guided microwave ablation with fine needle-assisted puncture positioning technique for hepatocellular carcinoma  
*Hao MZ, Hu YB, Chen QZ, Chen ZX, Lin HL*

1739  Clinicopathological characterization of ten patients with primary malignant melanoma of the esophagus and literature review  

1758  Endoscopic debulking resection with additive chemoradiotherapy: Optimal management of advanced inoperable esophageal squamous cell carcinoma  
*Ren LH, Zhu Y, Chen R, Shrestha Sachin M, Lu Q, Xie WH, Lu T, Wei XY, Shi RH*

1771  Nomogram for predicting the prognosis of tumor patients with sepsis after gastrointestinal surgery  
*Chen RX, Wu ZQ, Li ZY, Wang HZ, Ji JF*

1785  Efficacy and safety of laparoscopic radical resection following neoadjuvant therapy for pancreatic ductal adenocarcinoma: A retrospective study  
*He YG, Huang XB, Li YM, Li J, Peng XH, Huang W, Tang YC, Zheng L*

**Observational Study**

1798  To scope or not - the challenges of managing patients with positive fecal occult blood test after recent colonoscopy  
*Rattan N, Willmann L, Aston D, George S, Bassan M, Abi-Hanna D, Anandabaskaran S, Ermerak G, Ng W, Koo JH*

1808  Clinical implications of interleukins-31, 32, and 33 in gastric cancer  
*Liu QH, Zhang JW, Xia L, Wise SG, Hambly BD, Tao K, Bao SS*

1823  Construction and analysis of an ulcer risk prediction model after endoscopic submucosal dissection for early gastric cancer  
*Gong SD, Li H, Xie YB, Wang XH*

1833  Percutaneous insertion of a novel dedicated metal stent to treat malignant hilar biliary obstruction  
*Cortese F, Acquafredda F, Mardighian A, Zurlo MT, Ferraro V, Memeo R, Spiliopoulos S, Inchingolo R*

**EVIDENCE-BASED MEDICINE**

1844  Prediction of gastric cancer risk by a polygenic risk score of *Helicobacter pylori*  

**META-ANALYSIS**

1856  Dissecting novel mechanisms of hepatitis B virus related hepatocellular carcinoma using meta-analysis of public data  

1874  Prognostic and clinicopathological value of Twist expression in esophageal cancer: A meta-analysis  
*Song WP, Wang SY, Zhou SC, Wu DS, Xie JY, Liu TT, Wu XZ, Che GW*
LETTER TO THE EDITOR

1886 Nutrition deprivation affects the cytotoxic effect of CD8 T cells in hepatocellular carcinoma

Zhang CY, Liu S, Yang M
ABOUT COVER
Editorial Board Member of *World Journal of Gastrointestinal Oncology*, Luigi Marano, MD, PhD, Associate Professor, Department of Medicine, Surgery, and Neurosciences, University of Siena, Siena 53100, Italy. luigi.marano@unisi.it

AIMS AND SCOPE
The primary aim of *World Journal of Gastrointestinal Oncology* (WJGO, *World J Gastrointest Oncol*) is to provide scholars and readers from various fields of gastrointestinal oncology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJGO mainly publishes articles reporting research results and findings obtained in the field of gastrointestinal oncology and covering a wide range of topics including liver cell adenoma, gastric neoplasms, appendiceal neoplasms, biliary tract neoplasms, hepatocellular carcinoma, pancreatic carcinoma, cecal neoplasms, colonic neoplasms, colorectal neoplasms, duodenal neoplasms, esophageal neoplasms, gallbladder neoplasms, etc.

INDEXING/ABSTRACTING
The WJGO is now abstracted and indexed in PubMed, PubMed Central, Science Citation Index Expanded (SCIE, also known as SciSearch®), Journal Citation Reports/Science Edition, Scopus, Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2022 edition of Journal Citation Reports® cites the 2021 impact factor (IF) for WJGO as 3.404; IF without journal self cites: 3.357; 5-year IF: 3.250; Journal Citation Indicator: 0.53; Ranking: 162 among 245 journals in oncology; Quartile category: Q3; Ranking: 59 among 93 journals in gastroenterology and hepatology; and Quartile category: Q3. The WJGO’s CiteScore for 2021 is 3.6 and Scopus CiteScore rank 2021: Gastroenterology is 72/149; Oncology is 203/360.

RESPONSIBLE EDITORS FOR THIS ISSUE
Production Editor: Ying-Yi Yuan; Production Department Director: Xiang Li; Editorial Office Director: Jia-Ru Fan.
Prediction of gastric cancer risk by a polygenic risk score of *Helicobacter pylori*

Xiao-Yu Wang, Li-Li Wang, Shu-Zhen Liang, Chao Yang, Lin Xu, Meng-Chao Yu, Yi-Xuan Wang, Quan-Jiang Dong

**Abstract**

**BACKGROUND**
Genetic variants of *Helicobacter pylori* (*H. pylori*) are involved in gastric cancer occurrence. Single nucleotide polymorphisms (SNPs) of *H. pylori* that are associated with gastric cancer have been reported. The combined effect of *H. pylori* SNPs on the risk of gastric cancer remains unclear.

**AIM**
To assess the performance of a polygenic risk score (PRS) based on *H. pylori* SNPs in predicting the risk of gastric cancer.

**METHODS**
A total of 15 gastric cancer-associated *H. pylori* SNPs were selected. The associations between these SNPs and gastric cancer were further validated in 1022 global strains with publicly available genome sequences. The PRS model was established based on the validated SNPs. The performance of the PRS for predicting the risk of gastric cancer was assessed in global strains using quintiles and random forest (RF) methods. The variation in the performance of the PRS among different populations of *H. pylori* was further examined.

**RESULTS**
Analyses of the association between selected SNPs and gastric cancer in the global dataset revealed that the risk allele frequencies of six SNPs were significantly higher in gastric cancer cases than non-gastric cancer cases. The PRS model constructed subsequently with these validated SNPs produced significantly
higher scores in gastric cancer. The odds ratio (OR) value for gastric cancer gradually increased from the first to the fifth quintile of PRS, with the fifth quintile having an OR value as high as 9.76 (95% confidence interval: 5.84-16.29). The results of RF analyses indicated that the area under the curve (AUC) value for classifying gastric cancer and non-gastric cancer was 0.75, suggesting that the PRS based on *H. pylori* SNPs was capable of predicting the risk of gastric cancer. Assessing the performance of the PRS among different *H. pylori* populations demonstrated that it had good predictive power for cancer risk for hpEurope strains, with an AUC value of 0.78.

**CONCLUSION**
The PRS model based on *H. pylori* SNPs had a good performance for assessment of gastric cancer risk. It would be useful in the prediction of final consequences of the *H. pylori* infection and beneficial for the management of the infection in clinical settings.

**Key Words:** Polygenic risk scores; *Helicobacter pylori*; Gastric cancer; Single nucleotide polymorphism

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.

---

**Citation:** Wang XY, Wang LL, Liang SZ, Yang C, Xu L, Yu MC, Wang YX, Dong QJ. Prediction of gastric cancer risk by a polygenic risk score of *Helicobacter pylori*. *World J Gastrointest Oncol* 2022; 14(9): 1844-1855

**URL:** https://www.wjgnet.com/1948-5204/full/v14/i9/1844.htm

**DOI:** https://dx.doi.org/10.4251/wjgo.v14.i9.1844

---

**INTRODUCTION**

*Helicobacter pylori* (*H. pylori*) infection affects more than half of the world’s population[1,2]. The outcomes of *H. pylori* infection vary among individuals. The consequences of most infections are benign. However, a minority of infected individuals may eventually develop gastric cancer[3,4]. Predicting the outcomes of *H. pylori* infection is a major concern in the management of the infection. Substantial genetic variation has been found in the pathogen. Mutations cause increased virulence in certain strains, enhancing their carcinogenic potential[5,6]. It has been demonstrated that typing *H. pylori* strains based on the genetic variations of virulent genes has the potential to predict the risk of gastric cancer[7,8].

Two studies have been recently conducted to investigate the association between *H. pylori* genomic variations and gastric cancer within the hpEurope and hpEastAsia populations, respectively[9,10]. The first study contained 173 hpEurope strains and found 11 cancer risk-associated variants, including gene loss variants and single nucleotide polymorphisms (SNPs). Risk scores calculated based on the status of the *cag*11, *cag*12 and *cag*20 genes were increased during the progression from inflammation to gastric cancer. The other study identified 11 SNPs and three DNA motifs associated with gastric cancer through examination of 240 hpEastAsia strains. It is unclear whether the association between these variations and gastric cancer exists for all *H. pylori* strains. However, the findings from these studies suggest that SNPs from the *H. pylori* genome have the potential to predict the risk of gastric cancer.

To explore the combined effect of multiple SNPs on disease susceptibility, the polygenic risk score (PRS) model has been developed[11]. A PRS is calculated as a sum of the effects of multiple SNPs on disease. PRS models composed of SNPs from the human genome have been successfully used to predict the risk of cancers such as gastric cancer, colorectal cancer, and breast cancer[12-15]. Few studies, however, have been conducted to explore the capacities of PRS model constructed with SNPs from bacterial genomes in predicting the risk of cancer. Our study aimed to construct a PRS model based on validated risk alleles of *H. pylori* to predict the risk of gastric cancer.
MATERIALS AND METHODS

Strains and SNP selection
A total of 2022 H. pylori genome sequences deposited in GenBank at the National Center for Biotechnology Information by December 8, 2021 (https://www.ncbi.nlm.nih.gov/genome/browse#1/prokaryotes/169/), and the figshare website (https://figshare.com/s/2174da1fa20ae71c71e0)[10] were downloaded for further analyses. Of them, 1187 H. pylori strains had relevant clinical information of patients. Subsequently, duplicate strains and strains isolated from peptic ulcer disease, mucosa-associated lymphoma or stromal tumors were excluded from further analyses. This led to a final dataset of 1022 global strains included in the study. They were divided into gastric cancer (n = 253) and non-gastric cancer (n = 769) groups. Patients in the latter group were diagnosed with functional dyspepsia (n = 46), or chronic gastritis with or without intestinal metaplasia (n = 143 and n = 580, respectively). A total of 15 H. pylori SNPs or genetic variants from the two previous genome-wide association studies (GWASs) were selected for further analyses (Figure 1, Table 1)[9,10]. We cited high-quality articles in Reference Citation Analysis (https://www.referencitationanalysis.com). The selection criteria were as follows: (1) The length of the variants was no longer than five contiguous nucleotides; and (2) The SNP selected was located in a protein-coding region.

Construction of the neighbour-joining tree
Based on the 1022 H. pylori genomes, the SNPs in the core genome (present in > 99% isolates) were identified by aligning the assembled genomes against the reference genome (26695-1MET, accession number: CP010436.1) using MUMmer as previously described[16,17]. A neighbour-joining tree was then constructed based on the sequences of concatenated SNPs using TreeBeSt software (http://treesoft.sourceforge.net/treebest.shtml) with default parameters.

Statistical analyses
The chi-square test was used to test the difference in the prevalence of risk alleles in strains isolated from gastric cancer and non-gastric cancer. Student’s t test was used to compare the PRS values between the gastric cancer and non-cancer groups. These tests were performed using SPSS 18.0 software. Odds ratios (ORs) and 95% confidence intervals (CIs) of the selected SNPs were calculated using logistic regression analysis in R (version 3.6.3).

A PRS was created for each strain using the following equation: PRS = β1 + β2 + ... + βk + βn. Briefly, in this equation, βk is the value obtained from the regression analysis of the risk allele and disease, and n is the total number of SNPs included in the PRS[18]. Logistic regression analysis was performed to evaluate the association between PRS and gastric cancer risk and by quintiles of the PRS risk distribution, standardized by the controls, and using the 3rd quintile, 40%-60%, as the reference[18].

Random forest algorithm
A random forest (RF) model was built using the AUC-RF algorithm[19]. The input variables were the scores of each of the validated SNPs. A 20-times repeated 10-fold cross-validation of the RF model was performed. The performance of the RF model was demonstrated by receiver operating characteristic curve analysis[20].

RESULTS

Validation of SNPs in the global dataset
Previous studies have identified two sets of H. pylori SNPs that are associated with gastric cancer[9,10]. The association between these SNPs and gastric cancer has been verified only in strains from the hpEurope or hpEastAsia populations, respectively. We selected 15 SNPs to validate the association between selected SNPs and gastric cancer in global strains (Table 1). The risk alleles were defined as those with a higher prevalence in strains from gastric cancer. Statistical analyses revealed that the risk alleles of six SNPs showed a significant increase in prevalence in the gastric cancer group compared with the non-gastric cancer group. These SNPs, validated in the global dataset, were used for subsequent analyses.

Establishment of the six-SNP PRS model
To construct a PRS model for predicting the risk of gastric cancer, the logOR values of each validated SNP were calculated (Table 1). A PRS model was subsequently constructed with the sum of the logOR values of six validated SNPs. The mean PRS value was 8.64 ± 1.71 and 6.99 ± 1.27 in the gastric cancer and non-gastric cancer groups, respectively. The PRS value in the gastric cancer group was significantly higher (P = 5.6E-36).
Table 1 Selected and validated single nucleotide polymorphisms associated with gastric cancer in the global dataset

<table>
<thead>
<tr>
<th>SNP</th>
<th>Corresponding locus in the strain 26695</th>
<th>Gene name</th>
<th>Description</th>
<th>Position in gene</th>
<th>Position in chromosome</th>
<th>Risk allele</th>
<th>Amino acid change</th>
<th>Prevalence of risk allele</th>
<th>P value</th>
<th>LogOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1^</td>
<td>HP0082</td>
<td>tlpc</td>
<td>Chemotaxis sensor</td>
<td>163^</td>
<td>88029^</td>
<td>A^</td>
<td>K217E,Q</td>
<td>16.2%^</td>
<td>2.88E-15</td>
<td>1.29^</td>
</tr>
<tr>
<td>2^</td>
<td>HP0130</td>
<td>trit</td>
<td>BIR, Dps/NapA, RAD21 similarity</td>
<td>345^</td>
<td>140797^</td>
<td>C^</td>
<td>Synonymous</td>
<td>16.2%^</td>
<td>2.88E-15</td>
<td>2.26^</td>
</tr>
<tr>
<td>3^</td>
<td>HP0231</td>
<td>dstG/K</td>
<td>Thiol-disulfide interchange protein</td>
<td>433^</td>
<td>241625^</td>
<td>A^</td>
<td>T145A</td>
<td>23.2%^</td>
<td>5.34E-15</td>
<td>1.24^</td>
</tr>
<tr>
<td>4</td>
<td>HP0468</td>
<td>Unknown</td>
<td></td>
<td>729</td>
<td>489762</td>
<td>A</td>
<td>Synonymous</td>
<td>16.2%</td>
<td>2.88E-15</td>
<td>1.24^</td>
</tr>
<tr>
<td>5</td>
<td>HP0468</td>
<td>Unknown</td>
<td></td>
<td>705-708</td>
<td>489783-489786</td>
<td>CGCC</td>
<td>A236T</td>
<td>1.2%</td>
<td>0.110</td>
<td>2.36</td>
</tr>
<tr>
<td>6</td>
<td>HP0709</td>
<td>Adenosyl-chloride synthase</td>
<td></td>
<td>145</td>
<td>762953</td>
<td>A</td>
<td>N49D</td>
<td>14.6%</td>
<td>0.08</td>
<td>1.24^</td>
</tr>
<tr>
<td>7</td>
<td>HP0709</td>
<td>Adenosyl-chloride synthase</td>
<td></td>
<td>159</td>
<td>762967</td>
<td>A</td>
<td>Synonymous</td>
<td>90.0%</td>
<td>0.274</td>
<td>1.24^</td>
</tr>
<tr>
<td>8^</td>
<td>HP0747</td>
<td>trnB</td>
<td>tRNA ([guanine-N(7)-methyltransferase</td>
<td>(934-937)</td>
<td>(803467-803470)</td>
<td>GGAA</td>
<td>G312K,G,R+T313A,T,S</td>
<td>38.7%</td>
<td>1.17E-14</td>
<td>1.20^</td>
</tr>
<tr>
<td>9</td>
<td>HP0797</td>
<td>hpaA</td>
<td>Neuraminylfucose-binding hemagglutinin</td>
<td>334</td>
<td>854406</td>
<td>T</td>
<td>S112A</td>
<td>26.8%</td>
<td>0.567</td>
<td>1.24^</td>
</tr>
<tr>
<td>10^</td>
<td>HP0797</td>
<td>hpaA</td>
<td>Neuraminylfucose-binding hemagglutinin</td>
<td>325^</td>
<td>854415^</td>
<td>C^</td>
<td>L109F</td>
<td>40.1%^</td>
<td>1.12E-14</td>
<td>1.24^</td>
</tr>
<tr>
<td>11</td>
<td>HP0807</td>
<td>fecA-2</td>
<td>Iron importer in outer membrane</td>
<td>2010</td>
<td>861345</td>
<td>C</td>
<td>Synonymous</td>
<td>96.7%</td>
<td>0.158</td>
<td>1.24^</td>
</tr>
<tr>
<td>12^</td>
<td>HP1035</td>
<td>Outer membrane protein</td>
<td>798^</td>
<td>1117402^</td>
<td>A^</td>
<td>Synonymous</td>
<td>34.6%</td>
<td>2.93E-14</td>
<td>1.24^</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>HP1250</td>
<td>csd5</td>
<td>Cell shape determinant</td>
<td>370</td>
<td>1325727</td>
<td>A</td>
<td>N1161D</td>
<td>65.8%</td>
<td>1</td>
<td>1.24^</td>
</tr>
<tr>
<td>14</td>
<td>HP1440</td>
<td>isp</td>
<td>Inactive Ser protease</td>
<td>533</td>
<td>1513405</td>
<td>G</td>
<td>G173E</td>
<td>92.0%</td>
<td>0.505</td>
<td>1.24^</td>
</tr>
<tr>
<td>15</td>
<td>HP1467</td>
<td>ompA101</td>
<td>Outer membrane protein of OmpA family</td>
<td>53</td>
<td>1538114</td>
<td>T</td>
<td>II8T</td>
<td>98.4%</td>
<td>0.246</td>
<td>1.24^</td>
</tr>
</tbody>
</table>

^The single nucleotide polymorphisms validated in this work. Risk allele: Allele associated with gastric cancer. SNP: Single nucleotide polymorphism; OR: Odds ratio.

**Evaluation of the association between PRS and gastric cancer risk**

To evaluate the performance of the 6-SNP PRS model for predicting the risk of gastric cancer, the PRS values for each of the selected 1022 strains were grouped according to the quintile method. With the third quintile as the reference, the estimated OR value gradually increased from the first quintile (< 20%) to the fifth quintile (> 80%) (Figure 2, Table 2). The fifth quintile had an OR value as high as 9.76 (95% CI: 5.84-16.29).

To further confirm the combined effect of the validated SNPs for prediction of gastric cancer risk, an RF model was constructed with logOR values from each SNP as input. The classification potentials of the combined logOR values of validated SNPs were then analysed. The importance of each SNP is shown in Figure 3. The AUC value was 0.75 (DeLong 95% CI: 0.71-0.78), suggesting a good classifying capacity of the combined SNPs.

**Performance of risk prediction by PRS for different H. pylori populations**

Considering the remarkable genomic variations among strains from different H. pylori populations, the performance of PRS for predicting the risk of gastric cancer was subsequently assessed in different H. pylori populations. The results of the phylogenetic analyses divided the 1022 global strains into five groups, namely, the hpEastAsia, hpAsia2, hpEurope, America-related and Africa-related populations (Figure 4). Due to the small number of gastric cancer cases (2 cases in hpAsia2 and no cases in Africa-related populations), hpAsia2 and Africa-related populations were excluded from subsequent analyses.
Table 2 Associations between polygenic risk score and gastric cancer risk

<table>
<thead>
<tr>
<th>Quintile</th>
<th>Non-GC</th>
<th>GC</th>
<th>OR (95%CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>315 (41.0%)</td>
<td>44 (17.3%)</td>
<td>0.79 (0.46-1.34)</td>
<td>0.405</td>
</tr>
<tr>
<td>2</td>
<td>98 (12.7%)</td>
<td>7 (2.8%)</td>
<td>0.40 (0.17-0.97)</td>
<td>0.052</td>
</tr>
<tr>
<td>3</td>
<td>141 (18.3%)</td>
<td>25 (9.9%)</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>141 (18.3%)</td>
<td>49 (19.4%)</td>
<td>1.96 (1.15-3.35)</td>
<td>0.013</td>
</tr>
<tr>
<td>5</td>
<td>74 (9.6%)</td>
<td>128 (50.6%)</td>
<td>9.76 (5.84-16.29)</td>
<td>1.25E-14</td>
</tr>
</tbody>
</table>

CI: Confidence interval; OR: Odds ratio; GC: Gastric cancer

In analysing the performance of the established PRS model in different populations, the PRS value was higher in the gastric cancer group for all populations. Statistical analyses revealed a significant difference in PRS between the gastric cancer and non-gastric cancer groups in the hpEastAsia, hpEurope and America-related populations (Figure 5).

To further verify the combined effects of these SNPs for prediction of gastric cancer risk for different H. pylori populations, a RF classification model was built. The results of RF model analyses demonstrated that the AUC value was highest (0.78, DeLong 95%CI: 0.70-0.85) in the hpEurope population, suggesting a good ability of the combined SNPs to predict the risk of gastric cancer (Figure 6). However, the performance of the combined SNPs for risk prediction in other H. pylori populations was poor (Figure 6).

DISCUSSION

In this study, we constructed a PRS model based on validated H. pylori SNPs to predict the risk of gastric cancer. To our knowledge, our study is the first to evaluate a PRS model for cancer risk prediction constructed with genomic variants of H. pylori. H. pylori shows substantial genetic variations, resulting in remarkable interstrain differences in carcinogenic potential[5,21]. The presence/absence or large sequence variation of virulence genes and H. pylori SNPs have been shown to promote gastric carcinogenesis. Few studies have been conducted to assess the predictive power of these cancer-related genetic variations for gastric cancer[9,10]. Moreover, the combined effect of multiple variations on the predictive power for cancer risk has not been explored. Findings from this study demonstrate that a PRS model combining six H. pylori SNPs had a moderate capacity for prediction of gastric cancer risk. This is similar to the findings in studies on PRS model constructed with cancer-associated SNPs from the human genome[14,15].

To assess the combined effects of SNPs on gastric cancer risk prediction, we first selected 15 cancer-associated H. pylori SNPs from two previous GWAS studies. Their association has been validated in strains from specific geographical regions but not in a global strain collection. Our results demonstrated...
that only six of the SNPs showed a close association with gastric cancer in the global dataset. The SNPs at 88029, 241625, 803467 and 854415 in the reference strain 26695 caused nonsynonymous changes in the corresponding amino acid sequence, whereas the SNPs at 140797 and 1117402 in the reference strain 26695 produced synonymous variations. The \textit{hpaA} gene, harbouring the SNP at 854415, encodes an adhesion gene of \textit{H. pylori}\cite{22}. This gene is essential for colonization and is associated with the occurrence of gastric cancer\cite{23-25}. The SNP at 88029 was located on the \textit{tlpC} gene. \textit{TlpC} encodes a chemoreceptor that affects the chemotaxis of strains in the mouse gastric environment. It is associated with the induction of mucosal inflammation of the stomach\cite{26,27}. The SNP at 241625 was located in
dsbG/K, which has protein disulfide isomerase activity. DsbG/K interacts with a virulence-related factor in vitro[28,29]. In vitro studies have shown that a lack of dsbG/K may cause the loss of T4SS function and inhibit VacA secretion, which are considered the main pathogenic factors in H. pylori[30].

In this study, we constructed a PRS model with six SNPs validated in a global dataset. Assessments of the performance of the PRS model demonstrated that the PRS value was significantly higher in the gastric cancer group than in the non-gastric cancer group. A significant increase in the risk of gastric cancer was found across the quintiles of the PRS. These findings demonstrate that the six-SNPs PRS model is capable of predicting the risk of gastric cancer. In support of this finding, RF analyses demonstrated that the combination of the six SNPs has a high predictive power for gastric cancer, with an AUC value of 0.75. In a recent report, a PRS model constructed with SNPs from the human genome...
showed unsatisfactory power in classifying gastric cancer from healthy controls, with an AUC value of 0.56\cite{31}. It has been shown that a PRS model derived from 112 SNPs in the human genome and lifestyle factors possesses good predictive capacity for gastric cancer risk\cite{32}. For individuals infected with \textit{H. pylori}, assessment of their gastric cancer risk is of great concern in the clinical settings. Previous reports have demonstrated that certain genetic variants are associated with increased gastric cancer risk\cite{9,10}. Our study, for the first time, demonstrated the combined effect of \textit{H. pylori} genomic variations in the assessment of cancer risk. The PRS model derived from \textit{H. pylori} SNPs would have a high capacity in predicting gastric cancer risk for patients infected with the pathogen. This will benefit the clinical management of the prognosis of the \textit{H. pylori} infection. It is well known that age, gender and lifestyle factors, including alcohol consuming, smoking, diet habits and economic status, are closely associated with gastric cancer\cite{33-35}. In the future, a PRS model constructed with \textit{H. pylori} SNPs and those gastric cancer associated risk factors in this study would have substantially increased power in predicting the risk of gastric cancer. The \textit{H. pylori} genome shows great variations between strains\cite{36,37}. Genetic information differs greatly among \textit{H. pylori} populations, and their carcinogenic potential is also different \cite{5,21}. We thus evaluated the performance of the PRS model across \textit{H. pylori} populations. Our results demonstrated a good predictive power of PRS for hpEurope strains.

A limitation of this study is that the performance of the PRS model was not assessed in hpAsia2 and Africa-related \textit{H. pylori} populations because the number of strains with clinical information available was insufficient. Moreover, we could not consider age, gender, nutrition and other risk factors in the construction of the PRS model, as information on all of these risk factors was not consistently available across databases. A comprehensive risk model enclosing other risk factors of gastric cancer is indicated in future studies. Further \textit{in vitro} and \textit{in vivo} exploration of the roles of the combination of \textit{H. pylori} SNPs identified in this study in gastric cancer would be much helpful in supporting our findings.
CONCLUSION

In summary, we constructed a PRS model based on *H. pylori* SNPs, which showed great potential in the prediction of gastric cancer risk globally, especially for individuals infected with hpEurope strains. Findings from this study demonstrated that the PRS model constructed from bacteria genomic variations, in addition to the PRS model established with human SNPs, can be of great value for disease risk prediction. In clinical practice, it is usually difficult to assess gastric cancer risk in patients infected with *H. pylori*. The model constructed in this study would be beneficial for solving this issue.

ARTICLE HIGHLIGHTS

Research background
Multiple single nucleotide polymorphisms (SNPs) of *Helicobacter pylori* (*H. pylori*) associated with gastric cancer have been identified through bacterial genome-wide association studies. Polygenic risk score (PRS) calculated as a sum of effect of SNPs provides a tool for assessing genetic impact on diseases.

Research motivation
Predicting risk of gastric cancer is a major concern in the management of the *H. pylori* infection.

Research objectives
This study constructed a PRS model based on *H. pylori* SNPs to predict the risk of gastric cancer.

Research methods
Associations between previously reported *H. pylori* SNPs and gastric cancer were validated in global strains. The PRS model based on the validated SNPs was evaluated by quintiles and random forest (RF) methods.

Research results
A PRS model was constructed with six validated SNPs. Quintiles and RF methods demonstrated the combination of six SNPs has a high predictive power for gastric cancer.

Research conclusions
PRS model constructed from bacterial genomic variations can be of great value for gastric cancer risk prediction.

Research perspectives
Comprehensive risk models including personal and genomic information need to be established in future studies.

FOOTNOTES

Author contributions: Yang C and Liang SZ collected sequencing data; Xu L and Yu MC analyzed the data; Wang XY wrote the manuscript; Wang LL and Wang YX wrote the discussion part of the manuscript; Dong QJ designed the research and supervised the manuscript; and all authors reviewed the manuscript and approved the final version of the manuscript.

Supported by the National Natural Science Foundation of China, No. 31870777.

Conflict-of-interest statement: All the authors report no relevant conflicts of interest for this article.

PRISMA 2009 Checklist statement: The authors have read the PRISMA 2009 Checklist, and the manuscript was prepared and revised according to the PRISMA 2009 Checklist.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

Country/Territory of origin: China

ORCID number: Xiao-Yu Wang 0000-0002-3278-0879; Li-Li Wang 0000-0002-3607-0786; Chao Yang 0000-0003-0626-0586;
REFERENCES


S-Editor: Wang JJ
L-Editor: A
P-Editor: Wang JJ
H. pylori

Ge Z

and meta-analysis.

Quach DT

comprehensive review of population-based epidemiological studies.

Mihor A

wide association studies and prospective cohort study.

Chen Z, Hu Z, Wei Q, Shen H, Li L. Genetic risk, incident gastric cancer, and healthy lifestyle: a meta-analysis of genome-


Choi J

[PMID: 27138472]

Helicobacter pylori HP0231 (DsbK): role in disulfide bond formation, redox homeostasis and production of Helicobacter
cystein-rich protein HcpE.

Lester J

[PMID: 17517877]

Helicobacter pylori HP0231 Influences Bacterial Virulence and Is Essential for Gastric Colonization.

Epplein M

1993;

10.1128/jb.175.3.674-683.1993

19111390

1982; 143: 29-36

Yoon JY

[PMID: 21707676]

Recent mixing of Vibrio parahealmolyticus populations.

Yiangou K

[PMID: 29201240]

Combination of a 15-SNP Polygenic Risk Score and Classical Risk Factors for the Prediction of Breast Cancer Risk in


Carle ML, Urrea V, Boulestex AL, Malats N. AUC-RF: a new strategy for genomic profiling with random forest. Hum Hered
2011; 72: 121-132 DOI: 21996641 DOI: 10.1159/000330778

Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology
1982; 143: 274-7

Mishra V, Prasad MP, Chandra S, Chakraborty A, Mukherjee N, Aghi M, Deb A, Mukherjee S. Individualized Nutrition
for Gastric Cancer Patients: A Randomized Controlled Trial.


Wang XY et al. PRS of H. pylori

September 15, 2022 | Volume 14 | Issue 9 | WJGO | https://www.wjgnet.com 1854