









RESEARCH ARTICLE

Cancer Epidemiology

Infectious pathogens and risk of esophageal, gastric and duodenal cancers and ulcers in China: A case-cohort study

Christiana Kartsonaki^{1,2}  | Pang Yao¹ | Julia Butt³  | Rima Jeske³ | Catherine de Martel⁴  | Martyn Plummer⁵  | Dianjianyi Sun^{6,7} | Sarah Clark^{1,2} | Robin G. Walters^{1,2} | Yiping Chen^{1,2} | Jun Lv^{6,7}  | Canqing Yu^{6,7}  | Michael Hill^{1,2} | Richard Peto¹ | Liming Li^{6,7} | Tim Waterboer³ | Zhengming Chen^{1,2} | Iona Y. Millwood^{1,2}  | Ling Yang^{1,2} 

¹Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Oxford, UK

²Medical Research Council Population Health Research Unit (MRC PHRU), Nuffield Department of Population Health, University of Oxford, Oxford, UK

³Infections and Cancer Epidemiology Division, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁴Early Detection, Prevention and Infections Branch, International Agency for Research on Cancer, Lyon, France

⁵Department of Statistics, University of Warwick, Coventry, UK

⁶Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing, China

⁷Peking University Center for Public Health and Epidemic Preparedness and Response, Beijing, China

Correspondence

Iona Y. Millwood and Ling Yang, MRC PHRU & CTSU, Nuffield Department of Population Health, Big Data Institute Building, University of Oxford, Old Road Campus, Oxford, OX3 7LF, UK.
Email: iona.millwood@ndph.ox.ac.uk and ling.yang@ndph.ox.ac.uk

Abstract

Infection by certain pathogens is associated with cancer development. We conducted a case-cohort study of ~2500 incident cases of esophageal, gastric and duodenal cancer, and gastric and duodenal ulcer and a randomly selected subcohort of ~2000 individuals within the China Kadoorie Biobank study of >0.5 million adults. We used a bead-based multiplex serology assay to measure antibodies against 19 pathogens (total 43 antigens) in baseline plasma samples. Associations between pathogens and antigen-specific antibodies with risks of site-specific cancers and ulcers were assessed using Cox regression fitted using the Prentice pseudo-partial likelihood. Seroprevalence varied for different pathogens, from 0.7% for Hepatitis C virus (HCV) to 99.8% for Epstein-Barr virus (EBV) in the subcohort. Compared to participants seronegative for the corresponding pathogen, *Helicobacter pylori* seropositivity was associated with a higher risk of non-cardia (adjusted hazard ratio [HR] 2.73 [95% CI: 2.09–3.58]) and cardia (1.67 [1.18–2.38]) gastric cancer and duodenal ulcer (2.71 [1.79–4.08]). HCV was associated with a higher risk of duodenal cancer (6.23 [1.52–25.62]) and Hepatitis B virus was associated with higher risk of duodenal ulcer (1.46 [1.04–2.05]). There were some associations of antibodies against some herpesviruses and human papillomaviruses with risks of gastrointestinal cancers and ulcers but these should be interpreted with caution. This first study of multiple pathogens with risk of gastrointestinal cancers and ulcers demonstrated that several pathogens are associated with risks of gastrointestinal cancers and ulcers. This will inform future investigations into the role of infection in the etiology of these diseases.

Christiana Kartsonaki and Pang Yao are joint first authors.

Tim Waterboer, Zhengming Chen, Iona Y. Millwood and Ling Yang are joint senior authors.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *International Journal of Cancer* published by John Wiley & Sons Ltd on behalf of UICC.

Funding information

Cancer Research UK, Grant/Award Numbers: C500/A16896, C16077/A29186; Cancer Research UK Population Research Committee Project Award, Grant/Award Number: C56488/A24504; Kadoorie Charitable Foundation in Hong Kong; Wellcome Grants to Oxford University, Grant/Award Numbers: 088158/Z/09/Z, 104085/Z/14/Z, 202922/Z/16/Z, 212946/Z/18/Z; National Natural Science Foundation of China, Grant/Award Numbers: 82192900, 82192901, 82192904; National Key Research and Development Program of China, Grant/Award Number: 2016YFC0900500; The UK Medical Research Council, Grant/Award Numbers: MC_U137686851, MC_UU_12026/2, MC_UU_00017/1; British Heart Foundation, Grant/Award Number: CH/1996001/9454

KEYWORDS

duodenal cancer, duodenal ulcer, esophageal cancer, gastric cancer, infectious pathogens

What's new?

Approximately 13% of cancers worldwide have been attributed to infections with pathogens including *Helicobacter pylori* (*H. pylori*), human papillomaviruses and hepatitis virus B and C. However, prospective studies on the association of various infectious pathogens with risks of certain types of gastrointestinal cancers are scarce. In this large prospective cohort study investigating the role of 19 infectious pathogens and 43 antigens in the risk of gastrointestinal cancers among Chinese adults, *H. pylori* seropositivity was associated with a higher risk of non-cardia and cardia gastric cancer and duodenal but not gastric ulcer. HCV was associated with duodenal cancer and HBV with duodenal ulcer.

1 | INTRODUCTION

Chronic infection with certain viruses and bacteria play an important role in the etiology of cancer, causing about 2.2 million new cancer cases in 2018, with low- and middle-income countries such as China particularly affected.^{1–4} *Helicobacter pylori* (*H. pylori*), Human Papillomaviruses (HPV), Hepatitis B and C viruses (HBV, HCV) and Epstein–Barr virus (EBV) are the most important known cancer-causing pathogens, accounting for substantial proportions of gastric, cervical and liver cancer, and lymphoma worldwide, respectively.⁴ *H. pylori* infection is a well-established cause of peptic ulcer and gastric cancer,⁵ responsible for >300,000 cases of gastric cancer in China in 2018.⁶ There is also epidemiological evidence that HBV and EBV are associated with gastric cancer.^{7,8}

However, there are not many prospective studies on the associations of various infectious pathogens with risks of certain types of gastrointestinal cancers, or ulcers, especially in China where there are high incidence rates of infection-related cancers.⁴ Although the association of *H. pylori* infection with gastric cancer has been well-studied,^{5,6,9} previous studies have not investigated comprehensively the relationships between other pathogens and cancers at different sites along the gastrointestinal tract, as well as ulcers. Moreover, some studies had a retrospective design which may affect seropositivity to certain pathogens.

The aim of our study was to assess the associations between seropositivity for 19 infectious pathogens, measured using a multiplex serology assay, with risks of incident esophageal, gastric and duodenal cancers and ulcers in a case-subcohort study within a large cohort of Chinese adults.

2 | METHODS**2.1 | Study population and data collection**

The China Kadoorie Biobank (CKB) is a prospective cohort study of 512,715 Chinese adults.¹⁰ 210,205 men and 302,510 women aged

30–79 years were recruited into the study from 10 areas (5 urban and 5 rural) in China during 2004–2008.

At local study assessment clinics, participants completed an interviewer-administered laptop-based questionnaire on sociodemographic characteristics, smoking, alcohol consumption, diet, tea drinking, physical activity, personal and family medical history and current medication. A range of physical measurements, (e.g., anthropometry, blood pressure and lung function) were recorded by trained technicians, using calibrated instruments with standard protocols. A 10-ml non-fasting (with the time since the participant last ate recorded) blood sample was collected from participants into an EDTA vacutainer (BD Hemogard™, USA).

2.2 | Mortality and morbidity follow-up

The vital status of each participant was determined periodically through the Disease Surveillance Points (DSP) system of China CDC,¹¹ supplemented by regular checks against local records. In addition, information about the occurrence of major diseases and any episodes of hospitalization was collected through linkage, via each participant's unique national identification number, with disease registries and national health insurance claims databases. All events were coded using International Classification of Diseases 10th Revision (ICD-10) by trained staff who were blinded to baseline information.

By 1 January 2017, 44,037 (8.6%) participants had died, 4781 (0.9%) were lost to follow-up and 27,903 (5.4%) had developed cancer, including 2507 (0.5%) esophageal cancer (ICD-10 C15), 3464 (0.7%) gastric cancer (C16, among which 535 were gastric cardia cancer [C16.0]) and 107 (0.02%) duodenal cancer (C17.0) cases; 2911 (0.6%) participants had developed gastric ulcers (K25) and 1154 (0.2%) had duodenal ulcers (K26). Systematic cancer validation and adjudication is ongoing, with retrieval of original medical records from hospitals for any reported cancer cases to confirm the cancer

diagnosis along with collection of detailed clinical diagnosis information, for example, cancer sub-site, histopathological subtype, stage and grade. Among the ~19,000 cancer cases that have been validated in CKB, the overall accuracy of gastric cancer diagnosis was 92% (85% were of adenocarcinoma subtype) and 96% for esophageal cancer diagnosis (88.5% were squamous cell carcinoma).

2.3 | Case-cohort study

A case-cohort study design was used for the present study (Figure S1). Among all gastric cancer cases recorded at least 2 years after the start of follow-up and up to 1 January 2017 with an available plasma sample and no history of cancer, we selected all 437 recorded cardia gastric cancer (CGC) cases and randomly selected 500 non-cardia gastric cancer (NCGC) cases from 762 validation-confirmed cases. We included all 27 individuals with a validated diagnosis of esophageal adenocarcinoma (EAC), 500 other esophageal cancer cases and all 70 reported duodenal cancer cases. 300 gastric ulcer and 200 duodenal ulcer cases were selected at random from 1900 gastric ulcer and 790 duodenal ulcer reported cases, respectively, that occurred at least 2 years after the start of follow-up and did not have any cancer diagnosis in the 2 years after ulcer diagnosis or any time before ulcer diagnosis. A subcohort of 2000 participants was sampled using simple random sampling from the “modified baseline” cohort (surviving individuals with no history of cancer 2 years after entering the study who had an available plasma sample and had genotyping data available as part of a random sample of the cohort which had been selected for genotyping).

2.4 | Multiplex assay

Stored baseline plasma samples of 3950 participants were assayed with a custom-designed multiplex serology assay using a Luminex bead-based method, as described previously.^{12,13} This semi-quantitative assay measures the median fluorescence intensity (MFI), which corresponds to the levels of antibodies against an antigen of interest in a plasma sample.

We measured antibody levels against 43 antigens from 19 pathogens (Table S1), including herpesviruses (herpes simplex viruses 1 [HSV-1] and 2 [HSV-2], varicella zoster virus [VZV], EBV, cytomegalovirus [CMV], human herpesviruses 6 [HHV-6] and 7 [HHV-7]), HBV and HCV, human papillomaviruses 16 (HPV-16) and 18 (HPV-18), human polyomaviruses (BK, JC, MCV), human immunodeficiency virus (HIV), human T lymphotropic virus type 1 (HTLV-1), *Chlamydia trachomatis* (*C. trachomatis*), *Toxoplasma gondii* (*T. gondii*) and *H. pylori*.¹³ These 19 pathogens were either established or potential risk factors for cancer, or cardiovascular or neurodegenerative diseases, or are of novel scientific interest. Antigen-specific cut-offs to define seropositivity were used as described previously,¹⁴ and were quality assured by a visual inflection point method.¹⁵ The definitions for seropositivity for each pathogen are shown in Table S1.

2.5 | Statistical analysis

Of the 3950 participants assayed, after exclusions for assay or sample issues ($n = 5$), 3945 individuals (1964 cases and 1986 subcohort members) remained in the main analyses. Histograms were plotted to visually inspect the distributions of antibody levels to each antigen. We calculated Kendall's correlations between MFI levels of antibodies and Pearson's correlations between the seropositivity of pathogens.

The associations between seropositivity for each pathogen and each antigen with risk of each type of cancer and ulcer were assessed using Cox proportional hazards models, fitted using the Prentice pseudo-partial likelihood.¹⁶ Models were adjusted for age (numeric), sex, region (10 regions) and educational attainment (six groups: no formal education, primary school, middle school, high school, technical school/college, university), and time in study was used as the time scale. Individuals were followed up until the first occurrence of each type of event and were censored if they died of other causes, or were lost to follow-up, or to 1 January 2017, whichever occurred earlier. In the analysis of each subtype of gastric and esophageal cancer, respectively, individuals were followed up until the first occurrence of any gastric or esophageal cancer, respectively, and other subtypes were censored. In the analysis of gastric and duodenal ulcer, individuals in the subcohort were censored at the time of any cancer diagnosis to match the selection criteria of the ulcer case groups, and individuals with a history of peptic ulcer at baseline were excluded. The analysis was repeated additionally adjusting for age squared, smoking status (three groups: never, ex-regular, occasional, current regular smoking), alcohol drinking (three groups: never regular, ex-regular and current regular drinking) and body mass index (BMI, numeric). Adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) for each outcome by sero-status for each pathogen (and each antigen), were estimated. The plausibility of the proportional hazards assumption was assessed using plots of scaled Schoenfeld residuals and the associated χ^2 tests.^{17,18}

Analysis was done using R version 4.1.1¹⁹ and packages “survival”²⁰ and “cckbplotr.”²¹

3 | RESULTS

3.1 | Characteristics of individuals in the case-subcohort study

Among the 3945 participants that were included in the main analysis, the mean age at baseline of subcohort participants was lower than that of the cancer or ulcer cases (51.8 [standard deviation (SD) 10.8] and 58.5 [9.7], respectively). 61.7% of subcohort participants and 34.1% of cases were female. The proportions of participants living in an urban area were similar among subcohort participants and cases (50.2% and 44.5%, respectively). Proportions with at least 6 years of education and other socioeconomic factors (having a private toilet or

fridge) varied between the different study arms. Subcohort participants and duodenal ulcer cases were more likely to have had at least 6 years of education than the other study arms. The prevalence of smoking and alcohol drinking among men varied by study arm and was highest among men with NCGC and with esophageal cancer (Table 1). Levels of adiposity were similar in subcohort participants and cases.

3.2 | Associations of pathogen seropositivity with cancers and ulcers

As shown in our previously published article, seropositivity in the subcohort varied by pathogen, from 0.1% for HIV and HTLV to 99.8% for EBV, with low correlations between markers of different pathogens.¹³ Because seroprevalence of HIV and HTLV was too low for associations with outcomes to be reliably estimated, these pathogens are not included in the results. The seroprevalence of HPV-16, HHV-6 and HHV-7 was higher in women than in men.

Seropositivity for HBV and HPV-16 increased with age, and urban residence was associated with a higher seroprevalence of *H. pylori*. There were no clear associations with ever-regular smoking or alcohol drinking.¹³

Compared to seronegative participants, those seropositive for HHV-7 and HPV-18 had a higher risk of esophageal cancer (adjusted HR 1.67 [95% CI: 1.04–2.70] and 1.96 [1.05–3.66], respectively) (Figure 1). *H. pylori* infection was associated with a higher risk of NCGC (2.73 [2.09–3.58]) and CGC (1.67 [1.18–2.38]). More detailed results on the associations between *H. pylori* and gastric cancer have been reported separately as part of a meta-analysis.⁹ There was no association of *H. pylori* with esophageal cancer, which is considered to largely consist of non-adenocarcinoma subtypes (0.97 [0.70–1.34]). There was no evidence of any association of seropositivity for any pathogen with risk of EAC, and precision of these estimates was low (Figure 2). HSV-1 was associated with a lower risk of duodenal cancer (0.27 [0.07–0.97]), and HCV was associated with a higher risk of duodenal cancer (6.23 [1.52–25.62]).

TABLE 1 Baseline characteristics of the subcohort and cases.

	Subcohort (n = 1986)	Cases (n = 1964)	EC (n = 498)	EAC (n = 27)	CGC (n = 436)	NCGC (n = 497)	DC (n = 70)	GU (n = 197)	DU (n = 297)
Mean age, years	51.8	58.5	60.7	62.4	61.1	59.0	56.9	54.2	52.9
Female	61.7	34.1	26.9	25.9	25.5	34.4	44.3	49.8	42.1
Urban region	50.2	44.5	25.3	44.4	36.7	70.4	58.6	30.0	60.9
Education ≥6 years	51.8	38.6	28.5	40.7	32.6	45.1	45.7	38.0	59.9
Household size	3.7	3.7	3.6	3.6	3.7	3.5	3.8	4.0	3.8
Private toilet	55.7	42.7	24.5	29.6	33.7	62.6	65.7	41.1	53.8
Fridge	57.9	50.7	40.6	48.1	45.6	66.0	70.0	40.4	61.4
Ever regular smoking									
Males	74.1	81.3	84.1	80.0	76.0	87.4	71.8	79.9	76.3
Females	3.6	4.8	4.5	0	0	4.1	3.2	8.1	7.2
Current regular alcohol intake									
Males	33.8	37.1	45.9	30.0	27.4	41.1	28.2	31.5	34.2
Females	2.8	0.9	0.7	0	0	1.2	3.2	1.4	0
Mean BMI, kg/m ²	23.8	23.5	23.1	24.0	24.0	23.6	24.4	23.2	23.3
Daily consumption									
Meat	31.8	25.7	17.9	25.9	16.1	35.0	28.6	24.2	42.1
Preserved vegetables	17.7	22.3	18.7	29.6	24.3	32.2	11.4	14.5	17.3
Fresh fruit	22.0	14.0	7.4	18.5	7.6	24.9	18.6	11.1	18.8
Blood transfusions	4.0	4.9	5.0	3.7	5.0	4.4	5.7	4.0	7.1
HBV surface antigen positive	3.2	2.7	2.0	0	3.2	3.9	1.5	1.7	3.1
History of peptic ulcer	3.8	6.8	4.0	3.7	5.3	6.2	2.9	9.1	16.8
Treatment of peptic ulcer ^a	15.8	25.4	15.0	0	30.4	25.8	0	33.3	18.2

Note: Numbers are percentages (within arm), unless otherwise specified.

Abbreviations: CGC, cardia gastric cancer; DC, duodenal cancer; DU, duodenal ulcer; EAC, esophageal adenocarcinoma; EC, esophageal cancer; GU, gastric ulcer; NCGC, noncardia gastric cancer.

^aPercentages receiving treatment for peptic ulcer at baseline among participants with a history of peptic ulcer.

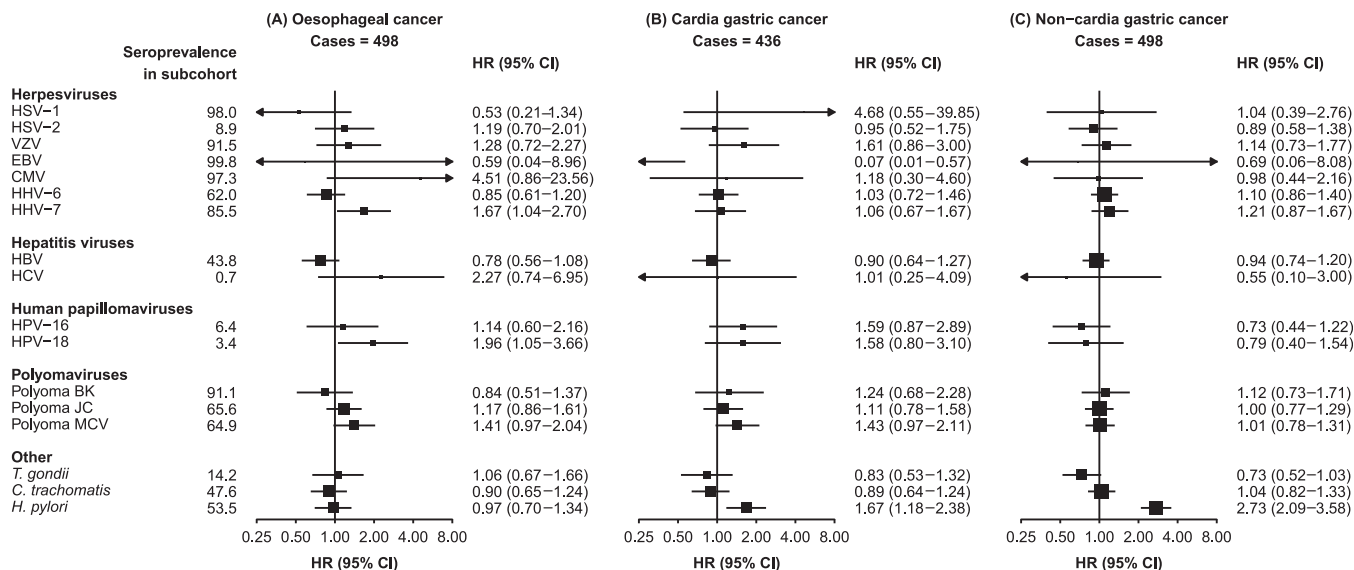


FIGURE 1 Adjusted HRs for risks of (A) esophageal cancer, (B) gastric cardia cancer and (C) non-cardia gastric cancer, associated with seropositivity to individual pathogens. HRs are adjusted for age, sex, region and education. Squares are the estimated HRs and line segments their 95% CIs. The area of the square is inversely proportional to the variance of the logHR.

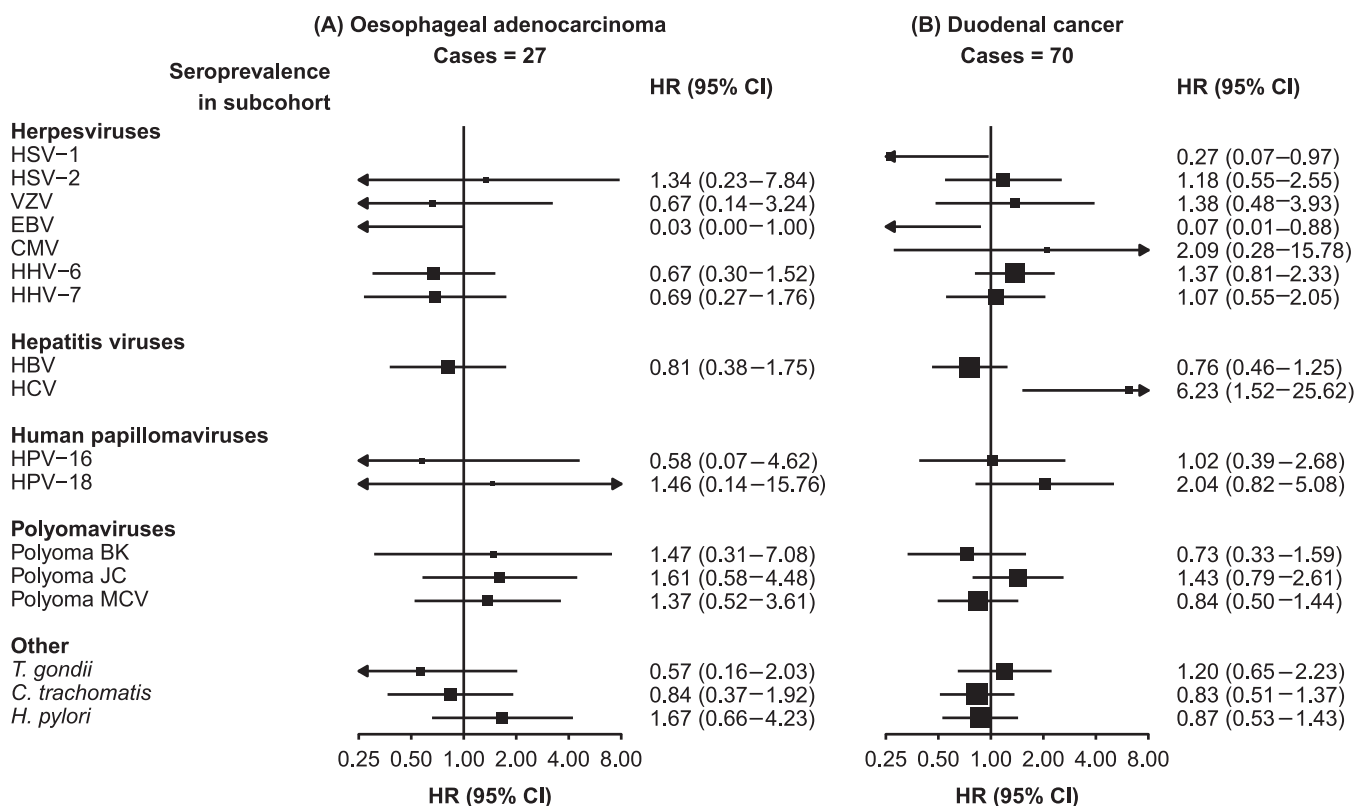


FIGURE 2 Adjusted HRs for risks of (A) esophageal adenocarcinoma and (B) duodenal cancer, associated with seropositivity to individual pathogens. HRs are adjusted for age, sex, region and education. Squares are the estimated HRs and line segments their 95% CIs. The area of the square is inversely proportional to the variance of the logHR.

Inverse associations of EBV with some types of cancer were observed, but the precision of the estimates was low due to the small number of EBV-negative individuals. There was no

evidence of any association between seropositivity for any of the polyomaviruses, *T. gondii* or *C. trachomatis* and any of the outcomes.

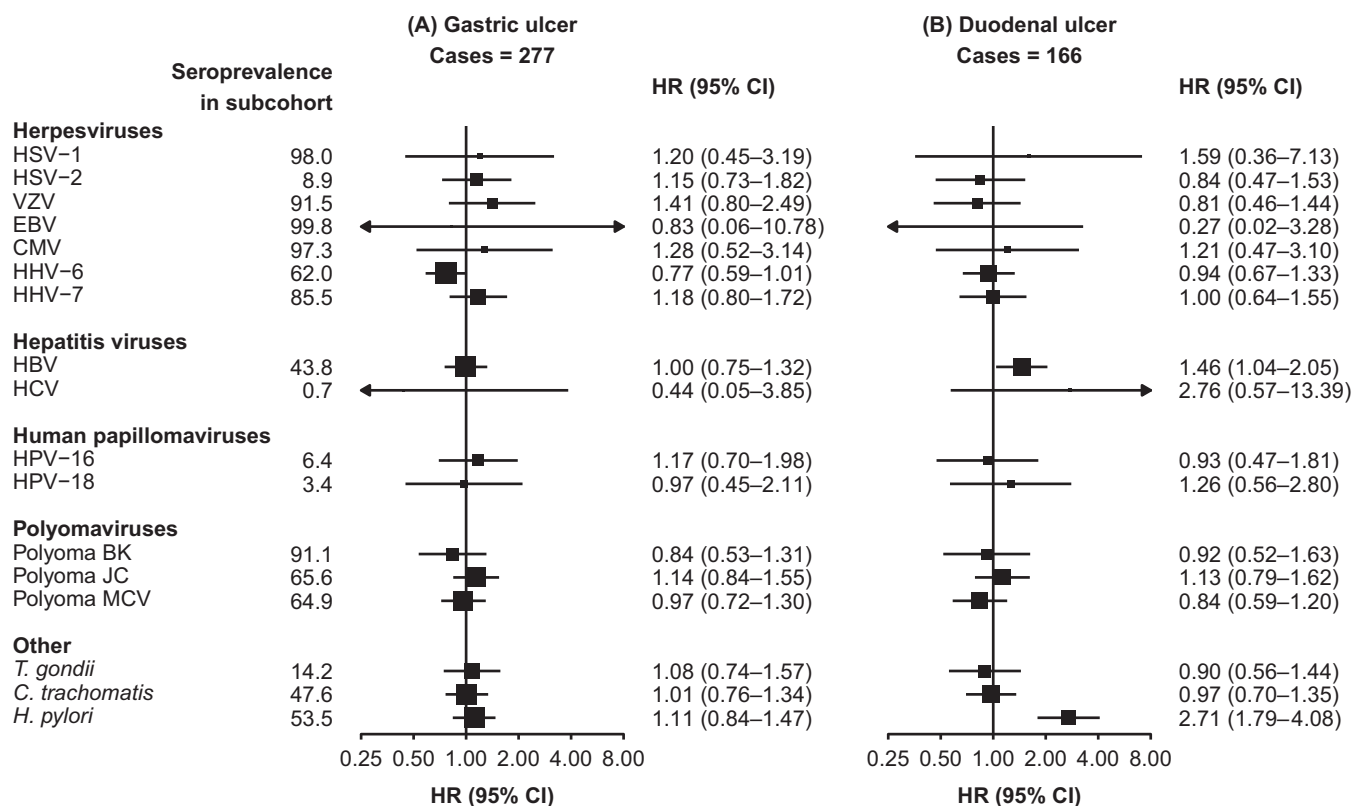


FIGURE 3 Adjusted HRs for risks of (A) gastric ulcer and (B) duodenal ulcer, associated with seropositivity to individual pathogens. Individuals with a history of peptic ulcer at baseline are excluded. HRs are adjusted for age, sex, region and education. Squares are the estimated HRs and line segments their 95% CIs. The area of the square is inversely proportional to the variance of the logHR.

HBV and *H. pylori* were associated with a higher risk of duodenal ulcer (1.46 [1.04–2.05] and 2.71 [1.79–4.08], respectively). However, none of the pathogens were associated with gastric ulcer (Figure 3).

3.3 | Associations of antigen seropositivity with cancers and ulcers

The seroprevalence of antibodies to *H. pylori* antigens in the subcohort varied from 11.4% (HpaA) to 70.8% (CagA). Several *H. pylori* antibodies were associated with CGC and NCGC (Figure S2) and duodenal ulcer (Figure S3). Seropositivity for HPV-16 E7 was associated with a higher risk of NCGC (1.43 [1.01–2.05]) (Figure S4). There was no evidence of any association between individual antibodies for pathogens other than *H. pylori* and risk of non-cardia gastric cancer (Figure S4). Seropositivity for EBV VCAp18 was inversely associated with risk of gastric cardia cancer (0.54 [0.32–0.92]). HCV NS3 was associated with a higher risk of duodenal cancer (4.20 [1.12–15.67]). HBc was associated with a higher risk of duodenal ulcer (1.45 [1.04–2.04]) (Figure S5). The magnitude of the association of HBe with duodenal ulcer was similar (1.40 [0.99–1.96]).

There was no evidence against the proportional hazards assumption in any of the analyses. Associations of pathogens (Figures S6 and S7) and antigens (Figures S8–S11) with cancers and ulcers did not change substantially with additional adjustment in sensitivity analyses.

4 | DISCUSSION

In this prospective study of multiple pathogens and risks of gastrointestinal tract cancers and ulcers among Chinese adults, we found that *H. pylori* seropositivity was associated with risks of developing NCGC, CGC and duodenal ulcer. Hepatitis viruses were associated with higher risks of duodenal ulcer and cancer, with a 46% higher risk of duodenal ulcer with HBV infection and a six-fold higher risk of duodenal cancer with HCV infection. HHV-7 seropositivity was associated with ~67% higher risk of esophageal cancer. No other pathogens were associated with the GI cancers and ulcers studied.

Although there was a lower seroprevalence of *H. pylori* in our study compared to prior studies in East Asians, the magnitude of its association with NCGC was remarkably consistent across studies in East Asians,^{14,15,22} and stronger than that in Western populations.²³

For example, *H. pylori* infection was associated with an odds ratio (OR) of 2.80 (2.25–3.48) for NCGC in a pooled analysis of eight prospective studies in East Asia (1608 cases),¹⁵ and with an OR of 1.90 (1.01–3.57) in the MCC-Spain multicenter case-control study (202 cases).²³ Previously published findings on the potential association between *H. pylori* and CGC have varied widely.^{9,23,24} To date, two case-control studies using multiplex serology have been conducted. *H. pylori* seropositivity was not associated with CGC in studies in Iran (142 cases; OR 1.70, 0.60–4.81)²⁴ or in Spain (61 cases; OR 0.54, 0.25–1.15).²³ Using the same multiplex serology assay, the present study found a significantly higher risk of CGC in Chinese with an *H. pylori* infection. This discrepancy may partly be due to the small number of cases included in the previous studies (CKB with >2-fold as many cases as in all the two studies combined), blood samples collected near cancer diagnosis and also be attributed to the fact that there are two distinct etiologies of CGC.²⁵ One type resembles NCGC, being a consequence of atrophic gastritis due to *H. pylori* infection and concentrating in East Asians. Another type resembles esophageal adenocarcinoma (EAC), is associated with gastroesophageal reflux and is mainly observed in Western populations. We did not observe a significantly higher risk of EAC associated with *H. pylori* infection; the estimates were, however, numerically similar to those for CGC, which may reflect this hypothesis of common pathophysiology between EAC and CGC. We found an association of *H. pylori* infection with duodenal ulcer, as expected, but not with gastric ulcer. This may be because *H. pylori* usually colonizes the antrum, which then leads to hypersecretion of acid, predisposing to duodenal ulcers.²⁶

The role of HBV infection in the development of hepatocellular carcinoma is well-established. As HBV infection also exists in gastric mucosa epithelial cells, it may be possible that HBV infection increases the risk of gastric cancer in a similar mechanism of HBV-related hepatocellular carcinoma,²⁷ or related to altered immune control. However, few epidemiological studies have shown conflicting results regarding the association between HBV infection and gastric cancer.²⁸ A previous study using the CKB showed that hepatitis B surface antigen (HBsAg) is associated with an HR of 1.41 (1.11–1.80) for stomach cancer (2157 cases), and the association was further replicated in two other small Chinese studies using polymerase chain reaction or immunohistochemistry test measurements of HBsAg.²⁹ In the present study, we measured HBV e antibody and core antibody levels and both were associated with higher risk of duodenal ulcer. HCV is also an established cause of hepatocellular carcinoma³⁰ and has been previously shown to be associated with a range of gastrointestinal and other cancers.³¹ The present study found that HCV infection was associated with a 6-fold higher risk of duodenal cancer, but the 95% confidence interval was wide due to the small number of cases and low prevalence of HCV. Certainly, the roles of HBV or HCV in gastric cancer development merit further investigations.

The oncogenic potential of HPV is well known in the context of cervical carcinoma, but its role in the development of esophageal cancer remains controversial. A recent meta-analysis involving

33 case-control studies worldwide suggested that HPV infection was associated with the risk of esophageal cancer (OR 1.62, 1.33–1.98).³² However, the mechanistic evidence using tumor tissue does not support the etiological role of HPV in esophageal carcinogenesis.³³ The E6 and E7 genes of the high-risk HPV types encode oncoproteins, and both act by interfering with the activity of cellular tumor suppressor proteins.³⁴ Epidemiological and mechanistic evidence on the causative role of HPV in NCGC is unclear. In the only previous prospective study involving only 70 NCGC cases, HPV-16 infection determined by ELISA was not associated with the risk of NCGC (OR 0.4, 0.1–1.6).³⁵ Overall, the present study is the first to explore the association of HPV and related oncogenic proteins with gastric cancer using multiplex serology measurements.

In our study, HHV-7 infection was reported for the first time to be associated with a higher risk of esophageal cancer. Several factors including immunosuppressive properties and pro-inflammatory properties indicated by their ability to alter the cytokine expression profile of infected cells suggest a possible role for HHV-7 in the pathogenesis of cancer.³⁶ However our findings should be interpreted with caution as multiple pathogens were assessed and some associations may be observed by chance. If a false discovery rate correction is considered, only the association of *H. pylori* with NCGC and duodenal ulcer would be considered significant.

A strength of the present study is its prospective design, with an exclusion of the first 2 years after blood collection to limit reverse causation, allowing estimation of associations between seropositivity of pathogens and future risk of disease. Moreover, we included well-characterized cases of different subtypes of cancers of adjacent sites and precursor lesions, which allows the assessment of relationships of infections with different stages of the underlying process of cancer development and allows comparison of associations with cancer subtypes to explore the extent to which there is common etiology. The use of a multiplex serology assay enabled the measurement of several antigens and the definition of seropositivity to several important pathogens. The availability of detailed participant characteristics allowed adjustment for potential confounders; however residual confounding may still exist, including from infections by other unmeasured pathogens. Another limitation of the study is the small number of cases for some disease subtypes, despite the large cohort from which they were identified, because of their low incidence in the population.

5 | CONCLUSION

In summary, in this prospective study of infections and risk of cancer and its precursors in China, we found associations between various pathogens and disease subtypes. In particular, HCV was associated with a higher risk of duodenal cancer, and *H. pylori* and several of its specific antigens were associated with CGC and NCGC and with duodenal ulcer. Further studies are needed to replicate these associations and to explore the potential mechanisms involved.

AUTHOR CONTRIBUTIONS

Christiana Kartsonaki: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – Original draft, Writing – Review and editing, Visualization, Funding acquisition; **Pang Yao:** Validation, Investigation, Writing – Original draft, Writing – Review and editing; **Julia Butt:** Methodology, Investigation, Resources, Data curation, Writing – Review and editing; **Rima Jeske:** Methodology, Investigation, Resources, Data curation, Writing – Review and editing; **Catherine de Martel:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – Review and editing, Funding acquisition; **Martyn Plummer:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – Review and editing, Funding acquisition; **Dianjianyi Sun:** Resources, Writing – Review and editing; **Sarah Clark:** Resources, Data curation, Writing – Review and editing; **Robin G. Walters:** Writing – Review and editing, Funding acquisition; **Yiping Chen:** Resources, Data curation, Writing – Review and editing; **Jun Lv:** Resources, Data curation, Writing – Review and editing; **Canqing Yu:** Resources, Data curation, Writing – Review and editing; **Michael Hill:** Methodology, Resources, Data curation, Writing – Review and editing; **Richard Peto:** Conceptualization, Investigation, Writing – Review and editing, Funding acquisition; **Liming Li:** Conceptualization, Investigation, Resources, Writing – Review and editing, Project administration, Funding acquisition; **Tim Waterboer:** Conceptualization, Methodology, Investigation, Resources, Writing – Review and editing, Project administration, Funding acquisition; **Zhengming Chen:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – Review and editing, Supervision, Project administration, Funding acquisition; **Iona Y. Millwood:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – Review and editing, Supervision, Project administration, Funding acquisition; **Ling Yang:** Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – Review and editing, Supervision, Project administration, Funding acquisition. The work reported in the article has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

The most important acknowledgement is to the participants in the study and the members of the survey teams in each of the 10 regional centers, as well as to the project development and management teams based at Beijing, Oxford and the 10 regional centers.

Members of the China Kadoorie Biobank Collaborative Group

International Steering Committee: Junshi Chen, Zhengming Chen (PI), Robert Clarke, Rory Collins, Liming Li (PI), Chen Wang, Jun Lv, Richard Peto, Robin Walters.

International Co-ordinating Centre, Oxford: Daniel Avery, Maxim Barnard, Derrick Bennett, Ruth Boxall, Ka Hung Chan, Yiping Chen, Zhengming Chen, Johnathan Clarke; Robert Clarke, Huaidong Du, Ahmed Edris Mohamed, Hannah Fry, Simon Gilbert, Pek Kei Im, Andri Iona, Maria Kakkoura, Christiana Kartsonaki, Hubert Lam, Kuang Lin, James Liu, Mohsen Mazidi, Iona Millwood, Sam Morris, Qunhua Nie, Alfred Pozarickij, Paul Ryder, Saredo Said, Dan Schmidt, Becky

Stevens, Iain Turnbull, Robin Walters, Baihan Wang, Lin Wang, Neil Wright, Ling Yang, Xiaoming Yang, Pang Yao.

National Co-ordinating Centre, Beijing: Xiao Han, Can Hou, Qingmei Xia, Chao Liu, Jun Lv, Pei Pei, Dianjianyi Sun, Canqing Yu.

10 Regional Co-ordinating Centres

Guangxi Provincial CDC: Naying Chen, Duo Liu, Zhenzhu Tang. **Liuzhou** CDC: Ningyu Chen, Qilian Jiang, Jian Lan, Mingqiang Li, Yun Liu, Fanwen Meng, Jinhuai Meng, Rong Pan, Yulu Qin, Ping Wang, Sisi Wang, Liuping Wei, Liyuan Zhou. **Gansu** Provincial CDC: Caixia Dong, Pengfei Ge, Xiaolan Ren. **Maiji** CDC: Zhongxiao Li, Enke Mao, Tao Wang, Hui Zhang, Xi Zhang. **Hainan** Provincial CDC: Jinyan Chen, Ximin Hu, Xiaohuan Wang. **Meilan** CDC: Zhendong Guo, Huimei Li, Yilei Li, Min Weng, Shukuan Wu. **Heilongjiang** Provincial CDC: Shichun Yan, Mingyuan Zou, Xue Zhou. **Nangang** CDC: Ziyang Guo, Quan Kang, Yanjie Li, Bo Yu, Qinai Xu. **Henan** Provincial CDC: Liang Chang, Lei Fan, Shixian Feng, Ding Zhang, Gang Zhou. **Huixian** CDC: Yulian Gao, Tianyou He, Pan He, Chen Hu, Huarong Sun, Xukui Zhang. **Hunan** Provincial CDC: Biyun Chen, Zhongxi Fu, Yuelong Huang, Huilin Liu, Qiaohua Xu, Li Yin. **Liuyang** CDC: Huajun Long, Xin Xu, Hao Zhang, Libo Zhang. **Jiangsu** Provincial CDC: Jian Su, Ran Tao, Ming Wu, Jie Yang, Jinyi Zhou, Yonglin Zhou. **Suzhou** CDC: Yihe Hu, Yujie Hua, Jianrong Jin, Fang Liu, Jingchao Liu, Yan Lu, Liangcai Ma, Aiyu Tang, Jun Zhang. **Qingdao** CDC: Liang Cheng, Ranran Du, Ruqin Gao, Feifei Li, Shanpeng Li, Yongmei Liu, Feng Ning, Zengchang Pang, Xiaohui Sun, Xiaocao Tian, Shaojie Wang, Yaoming Zhai, Hua Zhang, Licang CDC: Wei Hou, Silu Lv, Junzheng Wang. **Sichuan** Provincial CDC: Xiaofang Chen, Xianping Wu, Ningmei Zhang, Xiaoyu Chang. **Pengzhou** CDC: Xiaofang Chen, Jianguo Li, Jiaqiu Liu, Guojin Luo, Qiang Sun, Xunfu Zhong. **Zhejiang** Provincial CDC: Weiwei Gong, Ruying Hu, Hao Wang, Meng Wang, Min Yu. **Tongxiang** CDC: Lingli Chen, Qijun Gu, Dongxia Pan, Chunmei Wang, Kaixu Xie, Xiaoyi Zhang.

FUNDING INFORMATION

This work is funded by Cancer Research UK Population Research Committee Project Award (Ref: C56488/A24504). The CKB baseline survey and the first re-survey were supported by the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up has been supported by Wellcome Grants to Oxford University (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z) and grants from the National Natural Science Foundation of China (82192904, 82192901, 82192900) and from the National Key Research and Development Program of China (2016YFC0900500). The UK Medical Research Council (MC_UU_00017/1, MC_UU_12026/2 MC_U137686851), Cancer Research UK (C16077/A29186; C500/A16896) and the British Heart Foundation (CH/1996001/9454), provide core funding to the Clinical Trial Service Unit and Epidemiological Studies Unit at Oxford University for the project.

This research was funded in whole, or in part, by the Wellcome Trust [212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z]. For the purpose of Open Access, the author has applied a CC-BY public copyright license to any author accepted manuscript version arising from this submission.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The current study is based on the China Kadoorie Biobank (CKB). For open access data please visit <https://www.ckbiobank.org/data-access>. For the raw data from CKB that is utilized in this article contact ckbaccess@ndph.ox.ac.uk. Further information is available from the corresponding author upon request.

ETHICS STATEMENT

The China Kadoorie Biobank (CKB) complies with all the required ethical standards for medical research on human subjects. Ethical approvals were granted and have been maintained by the relevant institutional ethical research committees in the UK and China. Informed consent was obtained from all participants included in the study.

ORCID

Christiana Kartsonaki  <https://orcid.org/0000-0002-3981-3418>

Julia Butt  <https://orcid.org/0000-0001-7738-8611>

Catherine de Martel  <https://orcid.org/0000-0002-3642-277X>

Martyn Plummer  <https://orcid.org/0000-0001-5130-6497>

Jun Lv  <https://orcid.org/0000-0001-7916-3870>

Canqing Yu  <https://orcid.org/0000-0002-0019-0014>

Iona Y. Millwood  <https://orcid.org/0000-0002-0807-0682>

Ling Yang  <https://orcid.org/0000-0001-5750-6588>

REFERENCES

- Yim E-K, Park J-S. The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis. *Cancer Res Treat*. 2005;37(6):319-324. doi:10.4143/crt.2005.37.6.319
- Bouvard V, Baan R, Straif K, et al. WHO International Agency for Research on Cancer monograph working group. A review of human carcinogens—part B: biological agents. *Lancet Oncol*. 2009;10(4):321-322. doi:10.1016/s1470-2045(09)70096-8
- Arbyn M, Weiderpass E, Bruni L, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health*. 2020;8(2):e191-e203. doi:10.1016/S2214-109X(19)30482-6
- de Martel C, Georges D, Bray F, Ferlay J, Clifford GM. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. *Lancet Glob Health*. 2020;8(2):e180-e190. doi:10.1016/S2214-109X(19)30488-7
- Plummer M, Franceschi S, Vignat J, Forman D, de Martel C. Global burden of gastric cancer attributable to *Helicobacter pylori*. *Int J Cancer*. 2015;136(2):487-490. doi:10.1002/ijc.28999
- Yang L, Kartsonaki C, Yao P, et al. The relative and attributable risks of cardia and non-cardia gastric cancer associated with *Helicobacter pylori* infection in China: a case-cohort study. *Lancet Public Health*. 2021;6(12):e888-e896. doi:10.1016/S2468-2667(21)00164-X
- Tian T, Song C, Jiang L, et al. Hepatitis B virus infection and the risk of cancer among the Chinese population. *Int J Cancer*. 2020;147(11):3075-3084. doi:10.1002/ijc.33130
- Iizasa H, Nanbo A, Nishikawa J, Jinushi M, Yoshiyama H. Epstein-Barr virus (EBV)-associated gastric carcinoma. *Viruses*. 2012;4(12):3420-3439. doi:10.3390/v4123420
- Yao P, Kartsonaki C, Butt J, et al. *Helicobacter pylori* multiplex serology and risk of non-cardia and cardia gastric cancer: a case-cohort study and meta-analysis. *Int J Epidemiol*. 2023;52:dyad007. doi:10.1093/ije/dyad007
- Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol*. 2011;40(6):1652-1666. doi:10.1093/ije/dyr120
- Yang G, Rao C, Ma J, et al. Validation of verbal autopsy procedures for adult deaths in China. *Int J Epidemiol*. 2006;35(3):741-748. doi:10.1093/ije/dyi181
- Waterboer T, Sehr P, Michael KM, et al. Multiplex human papillomavirus serology based on in situ-purified glutathione s-transferase fusion proteins. *Clin Chem*. 2005;51(10):1845-1853. doi:10.1373/clinchem.2005.052381
- Yao P, Millwood I, Kartsonaki C, et al. Sero-prevalence of 19 infectious pathogens and associated factors among middle-aged and elderly Chinese adults: a cross-sectional study. *BMJ Open*. 2022;12(5):e058353. doi:10.1136/bmjopen-2021-058353
- Epplein M, Zheng W, Xiang Y-B, et al. Prospective study of *Helicobacter pylori* biomarkers for gastric cancer risk among Chinese men. *Cancer Epidemiol Biomarkers Prev*. 2012;21(12):2185-2192. doi:10.1158/1055-9965.EPI-12-0792-T
- Cai H, Ye F, Michel A, et al. *Helicobacter pylori* blood biomarker for gastric cancer risk in East Asia. *Int J Epidemiol*. 2016;45(3):774-781. doi:10.1093/ije/dyw078
- Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika*. 1986;73(1):1-11. doi:10.1093/biomet/73.1.1
- Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982;69(1):239-241. doi:10.2307/2335876
- Xue X, Xie X, Gunter M, et al. Testing the proportional hazards assumption in case-cohort analysis. *BMC Med Res Methodol*. 2013;13(1):88. doi:10.1186/1471-2288-13-88
- R: The R Project for Statistical Computing. Accessed February 28, 2020. <https://www.r-project.org/>
- Therneau TM (until 2009), Lumley T (original S→R port and R maintainer). *Survival: Survival Analysis*, 2019. Accessed February 28, 2020. <https://CRAN.R-project.org/package=survival>
- Wright N. ckbplot: Create CKB Plots. Accessed November 11, 2020. <https://github.com/neilstats/ckbplot>
- Murphy G, Freedman ND, Michel A, et al. Prospective study of *Helicobacter pylori* antigens and gastric noncardia cancer risk in the nutrition intervention trial cohort. *Int J Cancer*. 2015;137(8):1938-1946. doi:10.1002/ijc.29543
- Fernández de Larrea-Baz N, Pérez-Gómez B, Michel A, et al. *Helicobacter pylori* serological biomarkers of gastric cancer risk in the MCC-Spain case-control study. *Cancer Epidemiol*. 2017;50:76-84. doi:10.1016/j.canep.2017.08.002
- Shakeri R, Malekzadeh R, Nasrollahzadeh D, et al. Pylori serology and risk of gastric cardia and noncardia adenocarcinomas. *Cancer Res*. 2015;75(22):4876-4883. doi:10.1158/0008-5472.CAN-15-0556
- Mukaisho K, Nakayama T, Hagiwara T, Hattori T, Sugihara H. Two distinct etiologies of gastric cardia adenocarcinoma: interactions among PH, *Helicobacter pylori*, and bile acids. *Front Microbiol*. 2015;6:6. doi:10.3389/fmicb.2015.00412
- Testerman TL, McGee DJ, Mobley HLT. Adherence and colonization. In: Mobley HL, Mendz GL, Hazell SL, eds. *Helicobacter Pylori: Physiology and Genetics*. ASM Press; 2001.
- Wei X-L, Qiu M-Z, Jin Y, et al. Hepatitis B virus infection is associated with gastric cancer in China: an endemic area of both diseases. *Br J Cancer*. 2015;112(7):1283-1290. doi:10.1038/bjc.2014.406
- Li M, Wu S, Luo H, et al. Serological and molecular characterization of hepatitis B virus infection in gastric cancer. *Front Cell Infect Microbiol*. 2022;12:894836. doi:10.3389/fcimb.2022.894836
- Song C, Lv J, Liu Y, et al. China Kadoorie Biobank collaborative group. Associations between hepatitis B virus infection and risk of all cancer

- types. *JAMA Netw Open*. 2019;2(6):e195718. doi:[10.1001/jamanetworkopen.2019.5718](https://doi.org/10.1001/jamanetworkopen.2019.5718)
30. de Oliveria Andrade LJ, D'Oliveira A, Melo RC, De Souza EC, Costa Silva CA, Paraná R. Association between hepatitis C and hepatocellular carcinoma. *J Glob Infect Dis*. 2009;1(1):33-37. doi:[10.4103/0974-777X.52979](https://doi.org/10.4103/0974-777X.52979)
 31. Liu B, Zhang Y, Li J, Zhang W. Hepatitis C virus and risk of extrahepatic malignancies: a case-control study. *Sci Rep*. 2019;9(1):19444. doi:[10.1038/s41598-019-55249-w](https://doi.org/10.1038/s41598-019-55249-w)
 32. Wang J, Zhao L, Yan H, et al. A meta-analysis and systematic review on the association between human papillomavirus (types 16 and 18) infection and esophageal cancer worldwide. *PLoS One*. 2016;11(7):e0159140. doi:[10.1371/journal.pone.0159140](https://doi.org/10.1371/journal.pone.0159140)
 33. Halec G, Schmitt M, Egger S, et al. Mucosal alpha-papillomaviruses are not associated with esophageal squamous cell carcinomas: lack of mechanistic evidence from South Africa, China and Iran and from a world-wide meta-analysis. *Int J Cancer*. 2016;139(1):85-98. doi:[10.1002/ijc.29911](https://doi.org/10.1002/ijc.29911)
 34. Thomas JT, Laimins LA. Human papillomavirus oncoproteins E6 and E7 independently abrogate the mitotic spindle checkpoint. *J Virol*. 1998;72(2):1131-1137. doi:[10.1128/JVI.72.2.1131-1137.1998](https://doi.org/10.1128/JVI.72.2.1131-1137.1998)
 35. Kamangar F, Qiao Y-L, Schiller JT, et al. Human papillomavirus serology and the risk of esophageal and gastric cancers: results from a cohort in a high-risk region in China. *Int J Cancer*. 2006;119(3):579-584. doi:[10.1002/ijc.21871](https://doi.org/10.1002/ijc.21871)
 36. Alibek K, Baiken Y, Kakpenova A, et al. Implication of human herpesviruses in oncogenesis through immune evasion and Suppression. *Infect Agent Cancer*. 2014;9(1):3. doi:[10.1186/1750-9378-9-3](https://doi.org/10.1186/1750-9378-9-3)

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kartsonaki C, Yao P, Butt J, et al. Infectious pathogens and risk of esophageal, gastric and duodenal cancers and ulcers in China: A case-cohort study. *Int J Cancer*. 2024;154(8):1423-1432. doi:[10.1002/ijc.34814](https://doi.org/10.1002/ijc.34814)