Supplementary Table 1. Evaluation of the demographic information of current study population with the seventh national census data of China.

Category	Demographic information of the seventh national census in 2021[a]	Demographic information of the current study population
Average family size	2.62	2.90[b]
Percentage of men	43.34%	51.24%
Percentage of women	56.66%	48.76%
Age (population percentage)		
0-14	17.95%	10.32%
15-59	63.35%	67.52%
>60	18.70%	22.16%
City living (population percentage)	63.89%	74.61%
Living Regions (population percentage)		
Eastern China	39.93%	37.02%
Central China	25.83%	19.48%
Western China	27.12%	35.52%
Northeast China	6.98%	7.98%

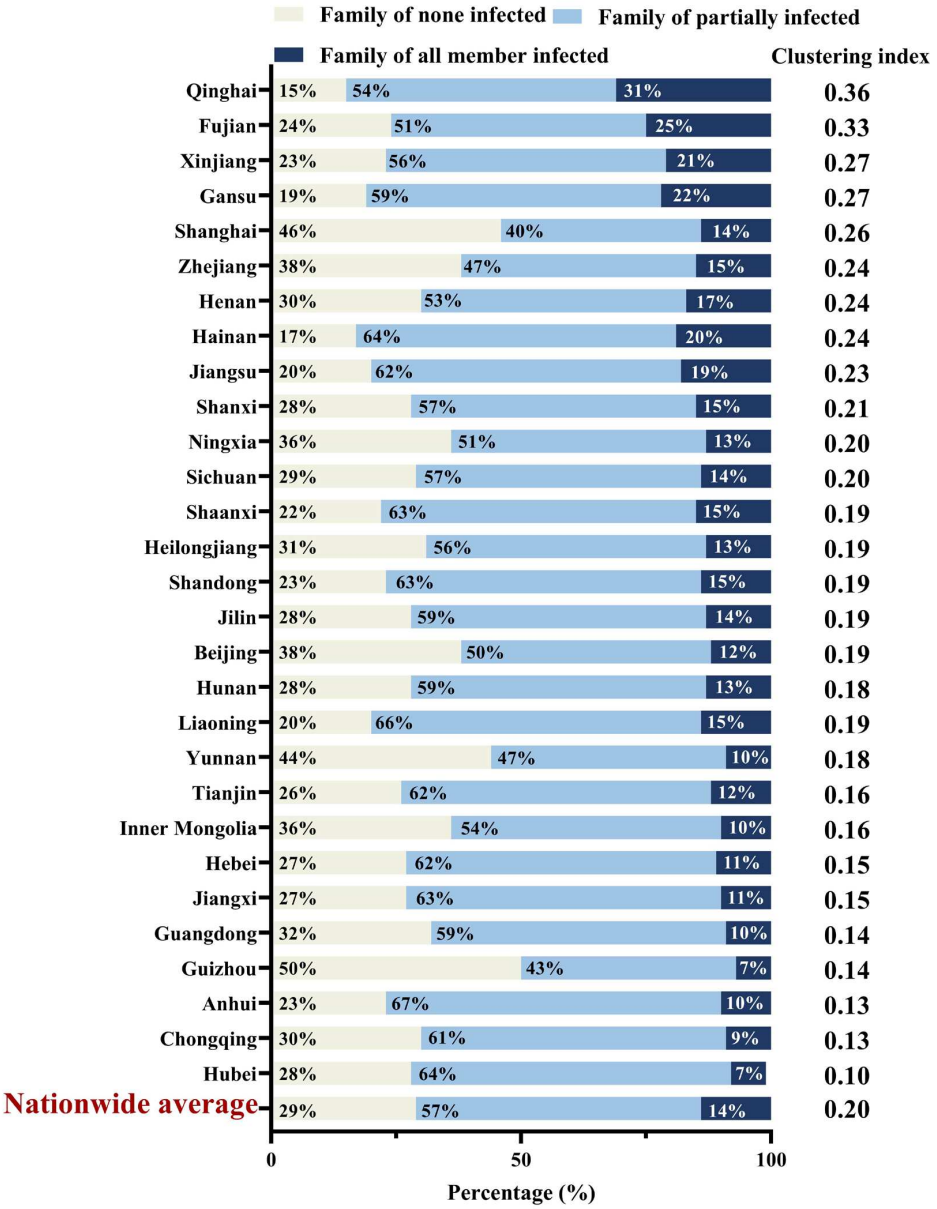
[a]Data of the seventh national census of China are from National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/tjgb/rkpcgb/). [b]Persons living alone was not enrolled as a household unit in the current study, which may explain the slightly larger average family size of the study population compared to the national census data.

Zhou X-Z, *et al*. *Gut* 2023; 72:855–869. doi: 10.1136/gutjnl-2022-328965

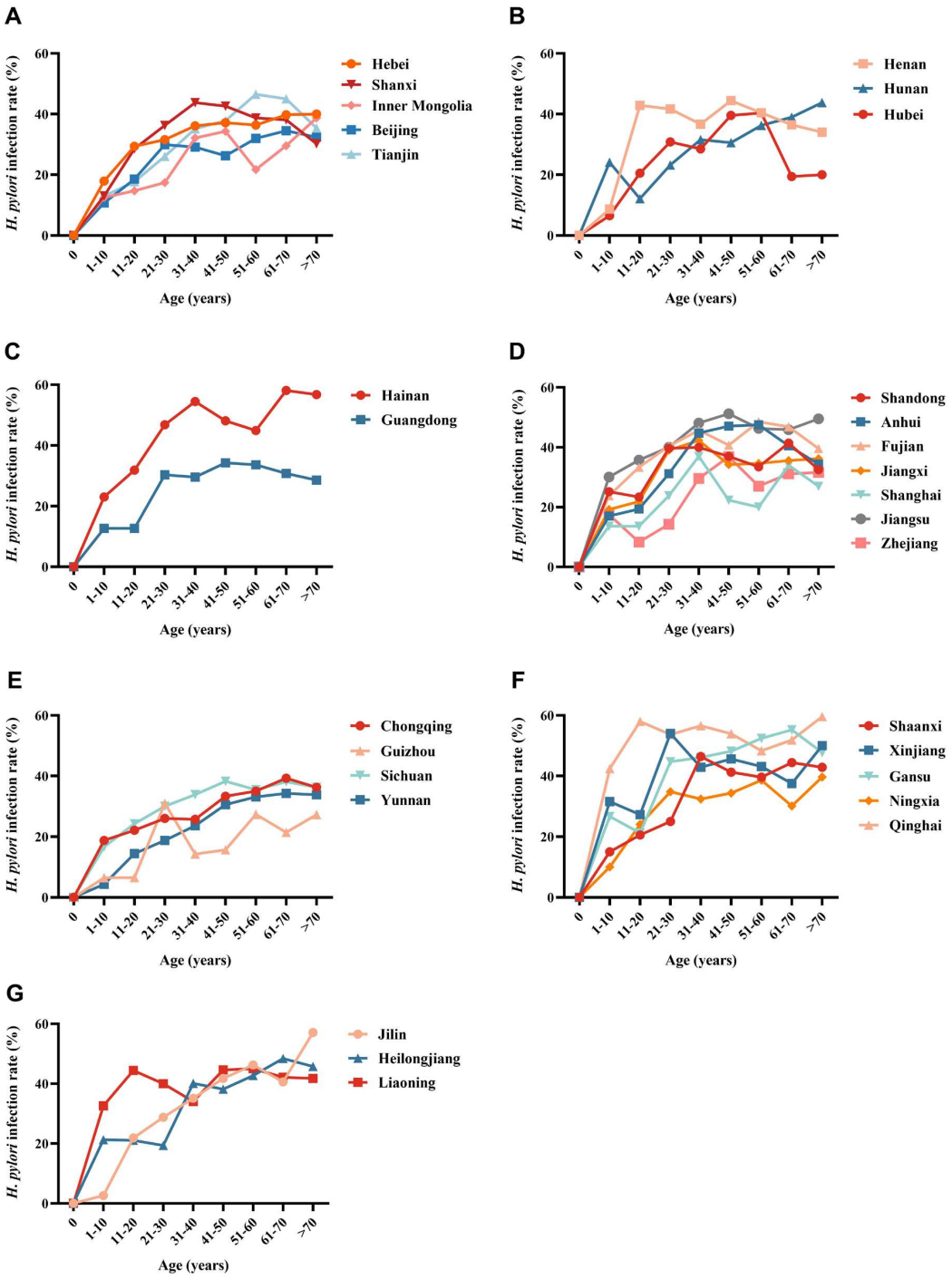
Supplementary Table 2. Correlation of family size and average number of infected person per families

Family size	Total families (n=10,735)					Infected families[a] (n=7636)				
	Number of families (%)	Average number of infected persons per family[b]	Percentage of infected persons[c] (%)	Average number of infected children and adolescents per family[d]	Percentage of infected children and adolescents[e](%)	Number of families (%)	Average number of infected persons per family[b]	Percentage of infected persons[c] (%)	Average number of infected children and adolescents per family[d]	Percentage of infected children and adolescents[e](%)
2	5305 (49.42)	0.84	41.83	0.02	2.07	3339 (43.73)	1.33	66.46	0.02	2.07
3	2909 (27.10)	1.22	40.72	0.08	6.87	2195 (28.75)	1.62	53.97	0.08	6.87
4	1461 (13.61)	1.55	38.72	0.14	8.93	1162 (15.22)	1.95	48.69	0.14	8.93
5	680 (6.33)	1.97	39.41	0.19	9.48	588 (7.70)	2.28	45.58	0.19	9.48
6	257 (2.39)	2.51	41.76	0.24	9.63	233 (3.05)	2.76	46.07	0.24	9.63
7 and above	123 (1.15)	3.31	47.27	0.28	8.60	119 (1.56)	3.42	43.11	0.28	8.60
Total	10,735 (100.00)	1.18	40.67	0.07	2.45	7636 (100.00)	1.66	54.53	0.10	3.29

[a]Infected family: at least one person in a family was infected by *H. pylori* (n=7636); n, sample number.
[b]Average number of infected persons per family: number of total infected persons in each subgroup / number of families in each subgroup.
[c]Percentage of infected persons: average number of infected persons per family[b]/average family size.
[d]Average number of infected children and adolescents per family: number of total infected children and adolescents in each subgroup / number of families in each subgroup.
[e]Percentage of infected children and adolescents: average number of infected children and adolescents per family[d]/average family size.

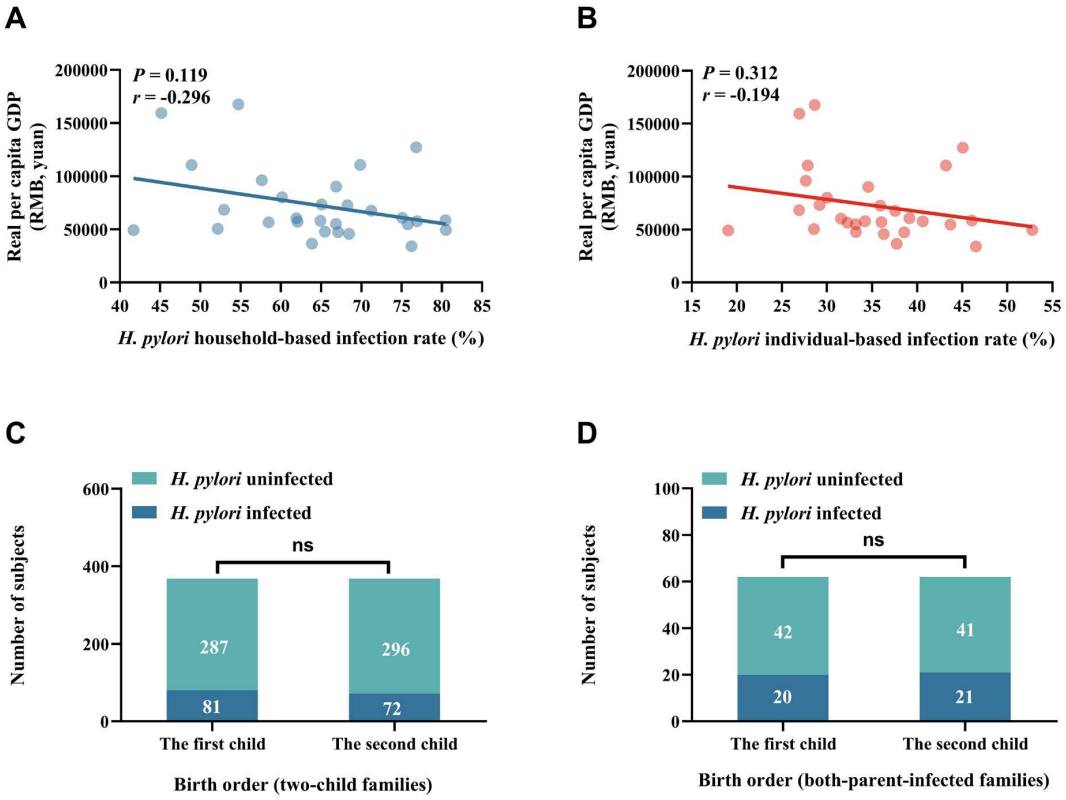


Supplementary Figure 1. The clustering index of *H. pylori* infection in 29 provinces of China. To understand *H. pylori* family-cluster infection among enrolled families, enrolled families were classified into three categories: family of none infected (grey), no family members were infected; family of partially infected (light blue), families had 1-5 infected members; and family of all member infected (dark blue), all members were infected within the household. Numbers within the bar indicate the percentage of each family subgroup.
Clustering level of *H. pylori*-infected family in each province was quantitatively assessed. Higher percentage indicated a higher degree of infection concentration within the particular province.



Supplementary Figure 2. *Helicobacter pylori* infection rate in 29 provinces by different geographic regions of China. Correlation of *H. pylori* infection with age in (A) North China, (B) Central China, (C) South China, (D) East China, (E) Southwest China, (F) Northwest China, and (G) Northeast China. The y axis represents *Helicobacter pylori* infection rate, and the x axis represents different age groups.

Zhou X-Z, *et al*. *Gut* 2023; 72:855–869. doi: 10.1136/gutjnl-2022-328965



Supplementary Figure 3. Correlation of *Helicobacter pylori* infection rate with gross income and birth order of children and adolescents within household.

(A) Association of *H. pylori* household-based infection rate and per capita gross domestic product (GDP) of 29 provinces. y axis represents the per capita GDP, and x axis represents household *H. pylori* infection rate (r=-0.296, P=0.119), each dot represents the data from one province.

(B) Association of *H. pylori* individual-based infection rate and per capita gross domestic product (GDP) of 29 provinces. y axis represents the per capita GDP, and x axis represents *H. pylori* infection rate (r=-0.194, P=0.312), each dot represents the data from one province.

(C) *H. pylori* infection rate of the first and second child in two-child families (P>0.05); and (D) *H. pylori* infection rate of the first and second child in both-parent-infected families. y axis represents children number, and x axis represents the birth order (P>0.05).

ns: not significant when comparing proportion of infected children between two groups (P>0.05).

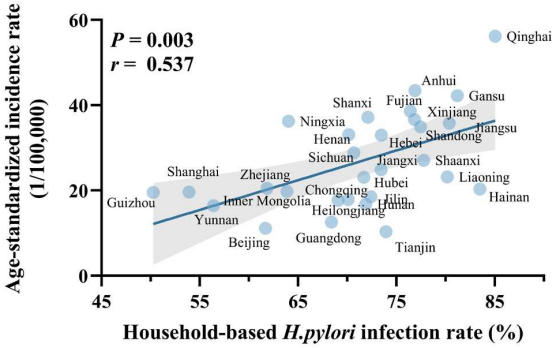
Child: participant's age is less than 18 years, including children and adolescents.

A

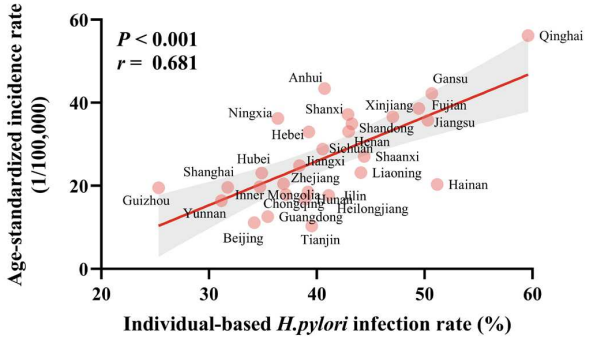
Gastric cancer incidence in 2013 and *H. pylori* infection rate in 2021 by province

No.	Provinces	Age-standardized incidence[a]	Household-based infection rate (%)[b]	Individua-based infection rate (%)[c]
1	Qinghai	56.16	85.06	59.61
2	Hainan	20.32	83.49	51.17
3	Gansu	42.24	81.19	50.67
4	Jiangsu	35.78	80.37	50.3
5	Liaoning	23.20	80.19	44.09
6	Shaanxi	27.05	77.78	44.4
7	Shandong	34.91	77.46	43.27
8	Anhui	43.42	76.89	40.72
9	Xinjiang	36.65	76.84	47.06
10	Fujian	38.62	76.42	49.46
11	Tianjin	10.32	73.95	39.52
12	Hebei	32.97	73.47	39.25
13	Jiangxi	24.86	73.45	38.39
14	Jilin	18.53	72.41	39.18
15	Shanxi	37.16	72.1	42.9
16	Hunan	16.75	71.88	38.91
17	Hubei	23.05	71.69	34.88
18	Sichuan	28.83	70.67	40.55
19	Henan	33.07	70.16	42.95
20	Chongqing	17.88	70.11	37.13
21	Heilongjiang	17.67	69.08	41.11
22	Guangdong	12.55	68.4	35.45
23	Ningxia	36.21	64.02	36.4
24	Inner Mongolia	19.78	63.87	34.71
25	Zhejiang	20.49	61.83	36.91
26	Beijing	11.13	61.69	34.2
27	Yunnan	16.45	56.4	31.16
28	Shanghai	19.61	53.92	31.73
29	Guizhou	19.52	50.27	25.33
30	Guangxi	25.12	NA	NA
31	Xizang	NA	NA	NA
32	Hong Kong	17.17	NA	NA
33	Macau	16.17	NA	NA
34	Taipei	37.12	NA	NA

B



C



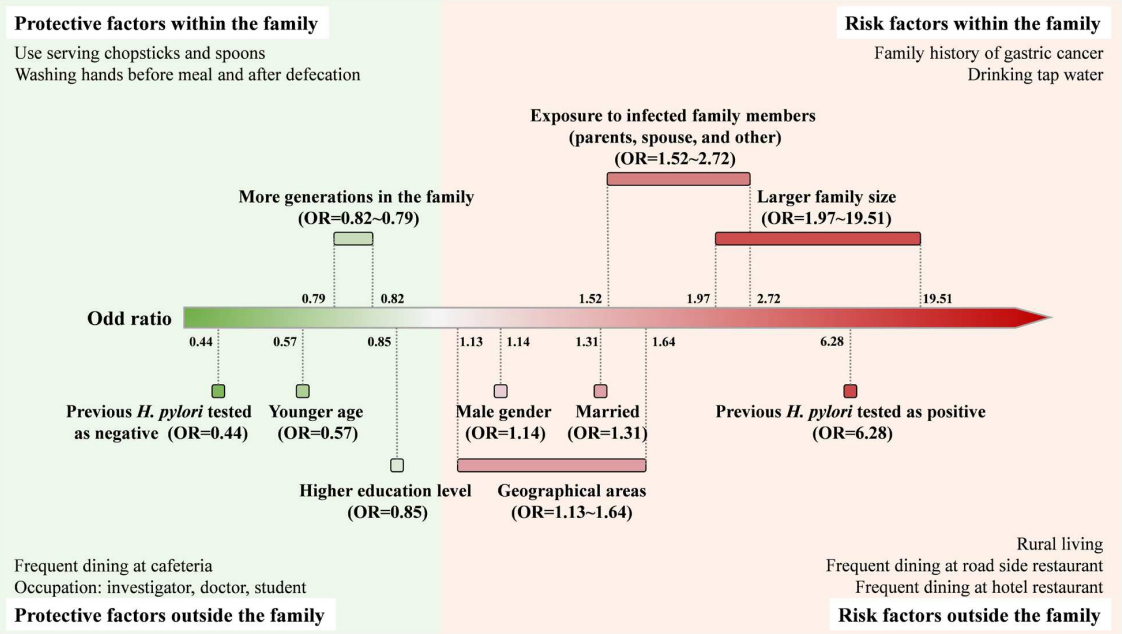
Supplementary Figure 4. Correlation of *Helicobacter pylori* infection with gastric cancer incidence in China.

(A) Age-standardized gastric cancer incidence rate in 2013 in all 31 provinces and 3 special regions in China[a], the data was applied and made into two new graphs with *H. pylori* infection data from current investigation (B, C).

(B) Correlation of household-based *H. pylori* infection rate and age-standardized incidence of gastric cancer in 29 provinces. y axis represents the age-standardized incidence of gastric cancer, and x axis represents household-based *H. pylori* infection rate (r=0.537, P=0.003), each dot represents the data from one province.

(C) Correlation of individual-based *H. pylori* infection rate and age-standardized incidence of gastric cancer in 29 provinces. y axis represents the age-standardized incidence of gastric cancer, and x axis represents individual-based *H. pylori* infection rate (r=0.681, P<0.001), each dot represents the data from one province.

[a] cite with permission from: *Wang BH, Wang N, Feng YJ, et al. Disease burden of stomach cancer in the Chinese population, in 1990 and 2013. Chinese Journal of Epidemiology, 2016, 37(6): 763-767).* [b]Household-based infection rate in 2021 is from the current study and is defined as percentage of infected household among all households. [c]Individual-based infection rate in 2021 is from the current study. NA, not available.



Supplementary Figure 5. Household *H. pylori* infection risk and protective factors

H. pylori infection risk and protective factor from within and outside the household were presented in two light purple (risk factor) and light green (protective factor) color areas respectively. They covered risk or protective factors from household-based, individual-based, and children and adolescent-based analysis. The upper parts of the figure are risk or protective factors within the household, while the external risk and protective factors are presented in the lower panel of the figure. Arrow indicated the odd ratio; red and green colored square blocks are the range of the odd ratio from independent risk or protective factors, while risk or protective factors are presented in the four corner of the figure.

Zhou X-Z, *et al*. *Gut* 2023; 72:855–869. doi: 10.1136/gutjnl-2022-328965