

**Supplementary Table 1** Univariate and multivariable analyses in the validation cohort

Variable	Validation cohort			P Value
	Total cohort (n=5091)	Gastric cancer (n=138)	Non-gastric cancer (n=4953)	
	No.	No. (%)	No. (%)	
Age, years				<0.001
40-49	1454	19 (13.8)	1435 (29)	
50-59	1703	35 (25.4)	1668 (33.7)	
60-69	1405	51 (37)	1354 (27.3)	
> 69	529	33 (23.9)	496 (10)	
Mean (SD)	56.3 (9.7)	62.1 (9.7)	56.1 (9.6)	< 0.001
Body mass index‡				
Mean (SD)	22.86 (2.88)	22.13 (3.03)	22.88 (2.87)	0.092
PG I, ng/ml				<0.001
≤ 41.92	312	26 (18.8)	286 (5.8)	
41.93-277.00	4474	100 (72.5)	4374 (88.3)	
> 277.00	305	12 (8.7)	293 (5.9)	
Median (range)	108.78 (921.58)	128.26 (531.68)	108.78 (921.58)	0.001
PG II, ng/ml				<0.001
≤ 2.78	265	6 (4.3)	259 (5.2)	
2.79-37.23	4539	122 (88.4)	4417 (89.2)	
> 37.23	287	10 (7.2)	277 (5.6)	
Median (range)	9.05 (263.10)	12.25 (93.88)	9 (263.10)	<0.001
PG I/II ratio				<0.001
≥ 3.89	4795	117 (84.8)	4678 (94.4)	
< 3.89	296	21 (15.2)	275 (5.6)	
Median (range)	11.84 (818.65)	8.71 (174.82)	11.95(818.65)	<0.001
G-17, pmol/l				
≤ 1.49	1411	17 (12.3)	1394 (28.1)	
1.50-5.70	1764	47 (34.1)	1717 (34.7)	
> 5.70	1916	74 (53.6)	1842 (37.2)	
Median (range)	3.44 (356.9)	6.38 (173.25)	3.40 (356.90)	<0.001
Sex				<0.001
Female	2557	39 (28.3)	2518 (50.8)	
Male	2534	99 (71.7)	2435 (49.2)	
Anti- <i>H. pylori</i> IgG, EIU				0.643
Negative (< 34)	2853	80 (58)	2773 (56)	
Positive (≥ 34)	2238	58 (42)	2180 (44)	
Family history				0.923
No	4413	120 (87)	4293 (86.7)	
Yes	678	18 (13)	660 (13.3)	
Smoking				0.031
No	3995	98 (71)	3897 (78.7)	
Yes	1096	40 (29)	1056 (21.3)	
Alcohol drinking				<0.001

No	4247	100 (72.5)	4147 (83.7)	
Yes	844	38 (27.5)	806 (16.3)	
High-salt diet				0.044
No	3078	72 (52.2)	3006 (60.7)	
Yes	2013	66 (47.8)	1947 (39.3)	
Pickled food				0.004
Occasional	4132	99 (71.7)	4033 (81.4)	
Regular	959	39 (28.3)	920 (18.6)	
Fried food				0.217
Occasional	4713	124 (89.9)	4589 (92.7)	
Regular	378	14 (10.1)	364 (7.3)	
Smoked food				0.136
Occasional	4767	125 (90.6)	4642 (93.7)	
Regular	324	13 (9.4)	311 (6.3)	
Barbecue food				0.192
Occasional	4812	127 (92)	4685 (94.6)	
Regular	279	11 (8)	268 (5.4)	
Overnight leftovers				0.119
Occasional	3622	90 (65.2)	3532 (71.3)	
Regular	1469	48 (34.8)	1421 (28.7)	
Red meat §				0.061
Occasional	2368	75 (54.3)	2293 (46.3)	
Regular	2723	63 (45.7)	2660 (53.7)	
White meat §				0.574
Occasional	2869	81 (58.7)	2788 (56.3)	
Regular	2222	57 (41.3)	2165 (43.7)	
Green vegetables				0.097
Occasional	1143	39 (28.3)	1104 (22.3)	
Regular	3948	99 (71.7)	3849 (77.7)	
Fresh fruits				0.003
Occasional	2428	83 (60.1)	2345 (47.3)	
Regular	2663	55 (39.9)	2608 (52.7)	

Data are presented as n (%) for categorical variables, or mean (SD) for continuous variables. For variables about eating habits, two categories for frequency of consumption were provided, that is, occasional (< 3 times/week) and regular (at least 3 times/week).

\* P values refer to comparison between gastric cancer and non-gastric cancer groups in the univariate analysis.

†For variables not significant ( $P > 0.05$ ) in the logistic regression model, multivariable data are not shown.

\*\*P values refer P values in the logistic regression multivariate analysis.

‡ Body mass index: weight (kg)/height (m)<sup>2</sup>.

§ Red meat includes beef, pork, and lamb. White meat includes fish, chicken, and duck.

CI, confidence interval; EIU, enzyme-immunoassay unit; G-17, gastrin-17; *H. pylori*, *Helicobacter pylori*; OR, odds ratio; PG, pepsinogen; SD, standard deviation.

**Supplementary Table 2** Multivariate analysis of the variables in the prediction model for junctional gastric cancer

Variable	Adjusted OR (95% CI)	<i>P</i> value
Age (years)		
50-59 vs. 40-49	1.20 (0.86-1.38)	0.763
60-69 vs. 40-49	4.67 (1.88-11.59)	<0.001
> 69 vs. 40-49	13.64 (5.53-33.64)	<0.001
Sex (Male vs. Female )	2.73 (1.28-5.84)	<0.001
Fried food (Regular vs. Occasional)	2.83 (1.16-6.91)	0.022
Fresh Fruits (Regular vs. Occasional)	0.34 (0.16-0.72)	0.005

CI, confidence interval; OR, odds ratio;

**Supplementary Table 3** Multivariate analysis of the variables in the prediction model for non-junctional gastric cancer

Variable	Adjusted OR (95% CI)	P value
Age (years)		
50-59 vs. 40-49	2.52 (1.55-4.07)	<0.001
60-69 vs. 40-49	3.64 (2.26-5.86)	<0.001
> 69 vs. 40-49	6.82 (4.12-11.28)	<0.001
Sex (Male vs. Female)	2.44 (1.83-3.26)	<0.001
PG I/II ratio (< 3.89 vs. $\geq$ 3.89)	2.19 (1.5-3.19)	<0.001
Anti- <i>H. pylori</i> (Positive vs. Negative)	1.33 (1.02-1.74)	0.036
G17 (pmol/L)		
1.50-5.70 vs. $\leq$ 1.49	2.08 (1.34-3.24)	0.001
>5.70 vs. $\leq$ 1.49	2.98 (1.96-4.54)	<0.001
Pickled food (Regular vs. Occasional)	1.50 (1.08-2.07)	0.015
Fried food (Regular vs. Occasional)	1.63 (1.06-2.49)	0.026

CI, confidence interval; OR, odds ratio;

**Supplementary Table 4** Multivariate analysis of the variables in the prediction model for intestinal gastric cancer

Variable	Adjusted OR (95% CI)	<i>P</i> value
Age (years)		
50-59 vs. 40-49	2.77 (1.72-4.47)	<0.001
60-69 vs. 40-49	4.13 (2.57-6.61)	<0.001
> 69 vs. 40-49	8.38 (5.13-13.68)	<0.001
Sex (Male vs. Female )	2.56 (1.95-3.37)	<0.001
PG I/II ratio (< 3.89 vs. ≥ 3.89)	1.98 (1.37-2.87)	0.003
Anti-H. pylori (Positive vs. Negative)	1.28 (1.00-1.64)	0.048
G17 (pmol/L)		
1.50-5.70 vs. ≤1.49	2.03 (1.35-3.07)	<0.001
>5.70 vs. ≤1.49	2.84 (1.92-4.2)	<0.001
Pickled food (Regular vs. Occasional)	1.53 (1.13-2.08)	0.006
Fried food (Regular vs. Occasional)	1.73 (1.16-2.58)	0.007

CI, confidence interval; OR, odds ratio;

**Supplementary Table 5** Multivariate analysis of the variables in the prediction model for diffuse gastric cancer

Variable	Adjusted OR (95% CI)	<i>P</i> value
Age (years)		
> 69 vs. 40-49	6.99 (1.56-31.26)	<0.001

CI, confidence interval; OR, odds ratio;

**Supplementary Table 6** Risk categories for different gastric diseases

		Gastric cancer	Low-grade intraepithelial neoplasia	Atrophic gastritis	Gastric ulcer	Duodenal ulcer	Gastric polyps	Reflux esophagitis
		267	545	2395	465	319	1250	369
Derivation cohort (n=9838)	Low risk	78 (29.2%)	336 (61.7%)	1538 (64.2%)	215 (46.2%)	205 (64.3%)	900 (72.0%)	228 (61.8%)
	Medium risk	120 (44.9%)	171 (31.4%)	709 (29.6%)	186 (40.0%)	94 (29.5%)	298 (23.8%)	123 (33.3%)
	High risk	69 (25.8%)	38 (7.0%)	148 (6.2%)	64 (13.8%)	20 (6.3%)	52 (4.2%)	18 (4.9%)
		138	281	1256	229	151	622	207
Validation cohort (n=5091)	Low risk	42 (30.4%)	155 (55.2%)	784 (62.4%)	102 (44.5%)	97 (64.2%)	446 (71.5%)	149 (72.0%)
	Medium risk	65 (47.1%)	100 (35.6%)	394 (31.4%)	104 (45.4%)	45 (29.8%)	153 (24.6%)	53 (25.6%)
	High risk	31 (22.5%)	26 (9.3%)	78 (6.2%)	23 (10.0%)	9 (6.0%)	24 (3.9%)	5 (2.4%)
Patients		405	826	3651	694	470	1872	576
Total(n=14929)	Low risk	120 (29.6%)	491 (59.4%)	2322 (63.6%)	317 (45.7%)	302 (64.2%)	1345 (71.8%)	377 (65.4%)
	Medium risk	185 (45.7%)	271 (32.8%)	1103 (30.2%)	290 (41.8%)	139 (29.6%)	451 (24.1%)	176 (30.6%)
	High risk	100 (24.7%)	64 (7.8%)	226 (6.2%)	87 (12.5%)	29 (6.2%)	76 (4.1%)	23 (4.0%)

**Supplementary Table 7** Multivariable analyses in the derivation cohort (excluding G-17)

Variable	Adjusted OR (95% CI)	<i>P</i> value
Age (years)		
50-59 vs. 40-49	2.85 (1.77-4.59)	< 0.001
60-69 vs. 40-49	4.43 (2.77-7.08)	< 0.001
> 69 vs. 40-49	9.40 (5.78-15.28)	< 0.001
Sex (Male vs. Female )	2.46 (1.88-3.22)	< 0.001
PG I/II ratio (< 3.89 vs. ≥ 3.89)	2.44 (1.71-3.48)	< 0.001
Anti-H. pylori (Positive vs. Negative)	1.37 (1.07-1.75)	0.014
Pickled food (Regular vs. Occasional)	1.48 (1.1-2.01)	0.011
Fried food (Regular vs. Occasional)	1.70 (1.15-2.53)	0.008

CI, confidence interval; OR, odds ratio;



**Supplementary Table 8** Performance of the risk prediction models with different sets of variables (with or without G-17)

	Prediction model	Predictors	Discrimination, AUC (95% CI)	Calibration, P value	P value
Models with G-17	Model 1	Age, sex, <i>H. pylori</i> infection, PG I/II ratio, G-17, pickled food, and fried food	0.760 (0.732-0.788)	0.605	Reference
	Model 2	Age, sex, <i>H. pylori</i> infection, PG I/II ratio, and G-17	0.752 (0.725-0.780)	0.629	0.689
	Model 3	Age, <i>H. pylori</i> infection, PG I/II ratio, and G-17	0.724 (0.696-0.753)	0.638	0.080
	Model 4	Sex, <i>H. pylori</i> infection, PG I/II ratio, and G-17	0.695 (0.666-0.723)	0.229	< 0.001
	Model 5	<i>H. pylori</i> infection, PG I/II ratio, and G-17	0.647 (0.615-0.678)	0.447	< 0.001
	Model 6	<i>H. pylori</i> infection and PG I/II ratio	0.569 (0.532-0.606)	0.834	< 0.001
	Model 7	PG I/II ratio and G-17	0.643 (0.611-0.675)	0.166	< 0.001
Models without G-17	Model 1	Age, sex, <i>H. pylori</i> infection, PG I/II ratio, pickled food, and fried food	0.738 (0.708-0.769)	0.860	Reference
	Model 2	Age, sex, <i>H. pylori</i> infection, PG I/II ratio	0.730 (0.700-0.760)	0.867	0.089
	Model 3	Age, <i>H. pylori</i> infection, PG I/II ratio	0.698 (0.668-0.728)	0.961	< 0.001
	Model 4	Sex, <i>H. pylori</i> infection, PG I/II ratio	0.643 (0.611-0.675)	0.900	< 0.001
	Model 5	<i>H. pylori</i> infection, PG I/II ratio	0.569 (0.535-0.602)	0.834	< 0.001
	Model 6	PG I/II ratio	0.546 (0.524-0.567)	No results	< 0.001

Model 1 included a set of variables selected according to statistical criteria. Other models included a subset of variables coming from the Model 1.

AUC, area under the receiver operating characteristic curve; CI, confidence interval; G-17, gastrin-17; *H. pylori*, *Helicobacter pylori*; PG, pepsinogen.

**Supplementary Table 9** Serum G-17 levels in patients with gastric cancer, atrophy, and other gastric diseases

G-17, pmol/L	Gastric cancer	Atrophy (n=2395)					Other gastric
	(n=267)	Total (n=2395)	Isolated corpus (n=48)	Angulus (n=71)	Antrum (n=579)	Several locations (n=1697)	diseases (n=7176)
≤ 1.49	34 (12.7%)	684 (28.6%)	7 (14.6%)	30 (42.3%)	183 (31.6%)	464 (27.4%)	2094 (29.2%)
1.50-5.70	83 (31.1%)	820 (34.2%)	11 (22.9%)	24 (33.8%)	211 (36.4%)	574 (33.8%)	2417 (33.7%)
> 5.70	150 (56.2%)	891 (37.2%)	30 (62.5%)	17 (23.9%)	185 (32.0%)	659 (38.8%)	2665 (37.1%)
Median (range)	6.889 (99.99)	3.528 (355.43)	8.55 (149.94)	2.01 (38.49)	3.13 (164.73)	3.6 (355.43)	3.44 (367.06)

**Supplementary Table 10** Distribution of histological types of GC in different locations.

		Junctional GC	Non-junctional GC
		36	231
Derivation cohort (n=267)	Intestinal type (n=167)	33	134
	Percentage in Row	19.8%	80.2%
	Percentage in Column	91.7%	58.0%
	Diffuse type (n=100)	3	97
	Percentage in Row	3%	97%
	Percentage in Column	8.3%	42.0%
		9	129
Validation cohort (n=138)	Intestinal type (n=93)	7	86
	Percentage in Row	7.5%	92.5%
	Percentage in Column	77.8%	67.7%
	Diffuse type (n=45)	2	43
	Percentage in Row	4.4%	95.6%
	Percentage in Column	22.2%	33.3%
		45	360
Total (n=405)	Intestinal type (n=260)	40	220
	Percentage in Row	15.4%	84.6%
	Percentage in Column	88.9%	61.1%
	Diffuse type (n=145)	5	140
	Percentage in Row	3.4%	96.6%
	Percentage in Column	11.1%	38.9%

**Supplementary Table 11** Risk of gastric cancer risk category between two groups  
in two cohorts

	Derivation cohort (n=267)		Validation cohort (n=138)	
	GC alone (n=179, %)	GC with other diseases (n=88, %)	GC alone (n=84, %)	GC with other diseases (n=54, %)
Low risk	57 (31.8%)	23 (26.1%)	28 (33.3%)	14 (25.9%)
Medium risk	82 (45.8%)	39 (44.3%)	38 (45.2%)	27 (50.0%)
High risk	40 (22.3%)	26 (29.5%)	18 (21.4%)	13 (24.1%)

**Supplementary Figure 1** Distribution of upper gastrointestinal diseases among all participants.