Table S1. Tests for multicollinearity in all variables.

Variable name	Variance inflation factor value
Serum 25(OH)D (ng/mL)	1.2
Race/Ethnicity	1.7
Sex	2.0
Age (years)	1.8
Family income to poverty ratio	1.3
Education level	1.4
Marital status	1.2
Place of birth	1.4
Food intake status	1.4
Smoking status	1.0
Alcohol consumption	1.3
Physical exercise	1.2
Source of drinking water	1.1
Peptic ulcer	1.1
Cardiovascular disease	1.1
Respiratory disease	1.0
Diabetes	1.1
Osteoporosis	1.0
Body mass index (kg/m^2)	1.1
Hemoglobin	1.7
Serum cholesterol	1.6
Serum triglycerides	1.6
Serum ferritin (ng/mL)	1.2
Serum folate (ng/mL)	1.4
Serum vitamin A (ug/dL)	1.6
Serum vitamin E (ug/dL)	2.3
Serum vitamin C (mg/dL)	1.0
Serum α carotene (ug/dL)	1.6
Serum β carotene (ug/dL)	1.6
Plasma glucose (mg/dL)	2.7
Glycated hemoglobin	2.7
Serum C-reactive protein (mg/dL)	1.1
Serum creatinine (mg/dL)	1.4
Serum thyroid stim hormone (U/L)	1.1
Serum thyroxine (ug/dL)	1.1

Note: Race/ethnicity (non-Hispanic white/black, Mexican-American, other); Sex(male, female); Family income to poverty ratio (<1.85, 1.85~3.5, >3.5); Education level (less than high school, high school, high school above); Marital status (married, never married, other); Place of birth (50 United States, Mexico, other); Food intake status (enough food to eat, sometimes not enough, often not enough); Smoking status (never, former smoker, current smoker); Alcohol consumption (never, < 1 time/month, 2~4 times/month, > 4 times/month); Physical exercise (no, yes); Source of drinking water (tapping water, bottled water); Peptic ulcer (no, yes); Cardiovascular disease includes congestive heart failure, hypertension, heart attack and stroke (no, yes); Respiratory disease includes asthma, chronic bronchitis and emphysema (no, yes); Diabetes (no, yes); Osteoporosis(no, yes); Body mass index (<25, 25~30, >30);

Table S2. Associations of covariates with *H. pylori* CagA seropositivity.

Covariate	*Beta*	*Se*	*Exp* (95% CI)	*p.* value
Family income to poverty ratio				
< 1.85	*Reference*	*Reference*	*Reference*	*Reference*
1.85~3.5	0.025	0.085	1.025 (0.868, 1.212)	0.769
> 3.5	-0.056	0.106	0.946 (0.768, 1.164)	0.598
Education level				
Less than high school	*Reference*	*Reference*	*Reference*	*Reference*
High school	-0.202	0.086	0.817 (0.690, 0.967)	0.019
Above high school	-0.005	0.109	0.995 (0.804, 1.232)	0.964
Marital status				
Never married	*Reference*	*Reference*	*Reference*	*Reference*
Married	0.088	0.129	1.092 (0.849, 1.404)	0.495
Other	0.032	0.093	1.033 (0.861, 1.238)	0.729
Place of birth				
50 United States	*Reference*	*Reference*	*Reference*	*Reference*
Mexico	0.009	0.120	1.009 (0.798, 1.277)	0.938
Other	0.430	0.175	1.537 (1.090, 2.168)	0.014
Smoking status				
Never	*Reference*	*Reference*	*Reference*	*Reference*
Former smoker	-0.083	0.101	0.921 (0.756, 1.122)	0.414
Current smoker	-0.007	0.087	0.993 (0.836, 1.178)	0.933
Alcohol consumption				
Never	*Reference*	*Reference*	*Reference*	*Reference*
< 1 time/month	0.412	0.173	1.511 (1.077, 2.119)	0.017
2~4 times/month	0.109	0.103	1.116 (0.912, 1.364)	0.287
> 4 times/month	0.069	0.095	1.071 (0.890, 1.290)	0.467
Physical exercise				
No	*Reference*	*Reference*	*Reference*	*Reference*
Yes	0.043	0.077	1.044 (0.898, 1.213)	0.579
Source of drinking water				
Tapping water	*Reference*	*Reference*	*Reference*	*Reference*
Bottled water	-0.031	0.122	0.970 (0.763, 1.232)	0.802
Peptic ulcer				
No	*Reference*	*Reference*	*Reference*	*Reference*
Yes	-0.006	0.081	0.995 (0.849,1.165)	0.945
Cardiovascular disease				
No	*Reference*	*Reference*	*Reference*	*Reference*
Yes	0.050	0.081	1.051 (0.896,1.232)	0.541
Respiratory disease				
No	*Reference*	*Reference*	*Reference*	*Reference*
Yes	-0.024	0.111	0.976 (0.785,1.214)	0.829
Diabetes				
No	*Reference*	*Reference*	*Reference*	*Reference*
Yes	0.035	0.123	1.036 (0.814,1.318)	0.773
Osteoporosis				
No	*Reference*	*Reference*	*Reference*	*Reference*
Yes	0.170	0.249	1.185 (0.728,1.930)	0.494
Body mass index (kg/m^2)				
< 25	*Reference*	*Reference*	*Reference*	*Reference*
25~30	-0.001	0.084	1.000 (0.849, 1.178)	0.996
> 30	0.007	0.095	1.007 (0.836, 1.213)	0.939
Hemoglobin	-0.060	0.030	0.942 (0.888, 0.999)	0.047
Serum cholesterol	-0.001	0.001	0.999 (0.998, 1.001)	0.281
Serum triglycerides	-0.001	0.001	1.000 (0.999, 1.000)	0.180
Serum ferritin (ng/mL)	-0.001	0.001	1.000 (0.999, 1.000)	0.041

Table S2. (Continued).

Covariate	*Beta*	*Se*	*Exp* (95% CI)	*p*. value
Serum folate (ng/mL)	-0.005	0.009	0.995 (0.978, 1.012)	0.535
Serum vitamin A (ug/dL)	-0.003	0.002	0.998 (0.993, 1.002)	0.270
Serum vitamin E (ug/dL)	-0.001	0.001	1.000 (0.999, 1.000)	0.107
Serum vitamin C (mg/dL)	-0.001	0.001	1.000 (0.998, 1.000)	0.347
Serum *α* carotene (ug/dL)	-0.002	0.008	0.998 (0.983, 1.014)	0.820
Serum *β* carotene (ug/dL)	0.003	0.002	1.003 (1.000, 1.006)	0.095
Plasma glucose (mg/dL)	-0.001	0.001	1.000 (0.998, 1.001)	0.867
Glycated hemoglobin	-0.028	0.032	0.973 (0.914, 1.036)	0.389
Serum C-reactive protein (mg/dL)	-0.047	0.046	0.954 (0.872, 1.044)	0.305
Serum creatinine (mg/dL)	0.003	0.117	1.003 (0.798, 1.261)	0.978
Serum thyroid stim hormone (U/L)	0.001	0.006	1.001 (0.990, 1.013)	0.857
Serum thyroxine (ug/dL)	0.002	0.018	1.002 (0.968, 1.037)	0.913

Note: Fixed variables (age, sex, race/ethnicity) were adjusted in the above analyses.

Table S3. Changes of the regression coefficient of 25(OH)D concentrations on *H. pylori* CagA seropositivity.

Covariate	Regression coefficient	
	Crude model [a] (Original regression coefficient = 0.0035)	**Complete model** [b] (Original regression coefficient = 0.0048)
Family income to poverty ratio	0.0035	0.0048
Education level	0.0030 *	0.0052
Marital status	0.0037	0.0045
Place of birth	0.0037	0.0044
Smoking status	0.0036	0.0047
Alcohol consumption	0.0037	0.0045
Physical exercise	0.0032	0.0051
Source of drinking water	0.0035	0.0048
Peptic ulcer	0.0035	0.0048
Cardiovascular disease	0.0034	0.0048
Respiratory disease	0.0035	0.0048
Diabetes	0.0035	0.0048
Osteoporosis	0.0035	0.0047
Body mass index (kg/m^2)	0.0035	0.0045
Hemoglobin	0.0035	0.0048
Serum cholesterol	0.0037	0.0048
Serum triglycerides	0.0032	0.0047
Serum ferritin (ng/mL)	0.0034	0.0048
Serum folate (ng/mL)	0.0042 *	0.0044
Serum vitamin A (ug/dL)	0.0044 *	0.0043 *
Serum vitamin E (ug/dL)	0.0044 *	0.0045
Serum vitamin C (mg/dL)	0.0035	0.0047
Serum α carotene (ug/dL)	0.0036	0.0045
Serum β carotene (ug/dL)	0.0022 *	0.0057 *
Plasma glucose (mg/dL)	0.0034	0.0046
Glycated hemoglobin	0.0033	0.0048
Serum C-reactive protein (mg/dL)	0.0035	0.0045
Serum creatinine (mg/dL)	0.0035	0.0048
Serum thyroid stim hormone (U/L)	0.0035	0.0047
Serum thyroxine (ug/dL)	0.0035	0.0049

Note: Fixed variables (age, sex, race/ethnicity) were adjusted in the above analyses. The change of the regression coefficient of 25(OH)D concentrations was observed by introducing the covariate in crude model or dropping the covariate in complete model.

[a] Crude model includes no covariates.

[b] Complete model includes all above covariates.

* The covariate changed the original regression coefficient more than 10%.

Table S4. Selected covariates as potential confounders in the study.

Dependent variable	Independent variable	Selected covariates (Standard 1) [a]	Selected covariates (Standard 2) [b]
H. pylori CagA seropositivity	25(OH)D concentrations	Education level Serum folate Serum vitamin A Serum vitamin E Serum β carotene	Education level Place of birth Alcohol consumption Hemoglobin Serum ferritin Serum folate Serum vitamin A Serum vitamin E Serum β carotene

[a] In Standard 1, covariates were selected by a change in effect estimate of more than 10% when they were introduced in crude model or dropped in complete model.
[b] In Standard 2, covariates further included those associated with *H. pylori* CagA seropositivity (defined as p-value < 0.1).

Table S5. Comparison of primary outcomes between complete data and imputed data from multiple imputation (adjusted OR and 95% CI for the association of serum 25(OH)D concentrations with *H. pylori* CagA seropositivity).

Effect modifier	Complete data [a]	Multiple imputation
Overall	1.01 (0.99,1.02)	1.02 (0.99,1.04)
Age Group (years)		
20~29	1.00 (0.97, 1.03)	1.00 (0.97, 1.03)
30~39	1.03 (0.99, 1.05)	1.03 (0.99, 1.06)
40~49	1.00 (0.97, 1.02)	0.99 (0.97, 1.03)
50~59	1.01 (0.98, 1.04)	1.01 (0.98, 1.04)
60~69	0.99 (0.97, 1.02)	0.99 (0.97, 1.02)
70~79	1.02 (0.99, 1.05)	1.02 (0.99, 1.05)
≥80	1.02 (0.98, 1.05)	1.02 (0.99, 1.05)
Sex		
Male	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Female	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Race/Ethnicity		
Non-Hispanic white	1.02 (1.00, 1.03) *	1.02 (1.00, 1.03) *
Non-Hispanic black	0.99 (0.97, 1.01)	0.99 (0.97, 1.02)
Mexican American	0.99 (0.98, 1.01)	0.99 (0.97, 1.01)
Other	1.08 (1.01, 1.15) *	1.08 (1.01, 1.16) *
Place of Birth		
50 United States	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Mexico	- [†]	- [†]
Other	1.09 (1.04, 1.14) *	1.09 (1.04, 1.15) *
Education Level		
Less than high school	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
High school	1.00 (0.98, 1.01)	1.00 (0.98, 1.01)
High school above	1.01 (0.99, 1.04)	1.02 (0.99, 1.04)
Alcohol Consumption		
Never	1.01 (0.99, 1.02)	1.01 (0.99, 1.02)
< 1 time/month	0.93 (0.88, 0.99) *	0.93 (0.88, 0.99) *
2~4 times/month	1.00 (0.98, 1.03)	1.00 (0.98, 1.03)
> 4 times/month	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Hemoglobin Level		
Quartile 1	1.01 (0.99, 1.03)	1.01 (0.99 ,1.03)
Quartile 2	1.00 (0.98, 1.02)	1.00 (0.98, 1.02)
Quartile 3	1.02 (0.99, 1.04)	1.02 (0.99, 1.04)
Quartile 4	1.03 (0.92, 1.17)	1.01 (0.99, 1.02)
Serum Ferritin Level		
Quartile 1	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Quartile 2	1.00 (0.98, 1.02)	1.00 (0.98, 1.02)
Quartile 3	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Quartile 4	1.00 (0.97, 1.02)	1.00 (0.98, 1.02)
Serum Folate Level		
Quartile 1	1.01 (0.99, 1.03)	1.02 (0.99, 1.04)
Quartile 2	1.00 (0.98, 1.02)	1.00 (0.98, 1.02)
Quartile 3	0.99 (0.97, 1.01)	0.99 (0.97, 1.01)
Quartile 4	1.02 (0.99, 1.04)	1.02 (0.99, 1.04)

Table S5. (Continued).

Effect modifier	Complete data [a]	Multiple imputation
Serum Vitamin A Level		
Quartile 1	1.00 (0.97, 1.02)	1.00 (0.97, 1.02)
Quartile 2	1.00 (0.98, 1.02)	1.00 (0.98 ,1.02)
Quartile 3	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Quartile 4	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Serum Vitamin E Level		
Quartile 1	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Quartile 2	1.00 (0.98, 1.02)	1.00 (0.98, 1.03)
Quartile 3	1.02 (0.99, 1.04)	1.02 (0.99, 1.04)
Quartile 4	0.99 (0.97, 1.01)	0.99 (0.97, 1.01)
Serum β-Carotene Level		
Quartile 1	1.00 (0.98, 1.02)	1.00 (0.98, 1.02)
Quartile 2	0.99 (0.98, 1.01)	0.99 (0.98, 1.01)
Quartile 3	1.01 (0.99, 1.03)	1.01 (0.99, 1.03)
Quartile 4	1.02 (0.99, 1.04)	1.02 (0.99, 1.04)

Note: Factors were not adjusted when they were used as effect modifiers.

[a] Complete data did not include missing data.

[†] The model failed because of the small sample size of other covariates in this stratum.

* Statistically significant associations ($p < 0.05$).

Table S6. Comparison of primary outcomes between complete data and imputed data from multiple imputation (adjusted OR and 95% CI for the association of vitamin D deficiency with *H. pylori* CagA seropositivity).

Effect modifier	Complete data	Multiple imputation
Overall	0.90 (0.77, 1.06)	0.90 (0.77, 1.05)
Age Group (years)		
20~29	0.99 (0.64, 1.52)	0.96 (0.63, 1.46)
30~39	0.81 (0.53, 1.25)	0.78 (0.52, 1.20)
40~49	0.94 (0.63, 1.41)	0.95 (0.64, 1.41)
50~59	0.83 (0.52, 1.34)	0.82 (0.51, 1.31)
60~69	0.83 (0.57, 1.23)	0.87 (0.60, 1.27)
70~79	0.97 (0.61, 1.54)	0.96 (0.61, 1.52)
⩾80	0.68 (0.41 1.12)	0.70 (0.43, 1.13)
Sex		
Male	1.03 (0.83, 1.30)	1.02 (0.82, 1.27)
Female	0.79 (0.62, 1.00)	0.80 (0.64, 1.00)
Race/Ethnicity		
Non-Hispanic white	0.69 (0.52, 0.92) *	0.69 (0.53, 0.92) *
Non-Hispanic black	1.17 (0.83, 1.65)	1.12 (0.80, 1.56)
Mexican American	1.01 (0.79, 1.30)	1.01 (0.79, 1.29)
Other	0.52 (0.20, 1.34)	0.57 (0.23, 1.43)
Place of Birth		
50 United States	0.98 (0.81, 1.80)	0.95 (0.79, 1.15)
Mexico	- †	- †
Other	0.43 (0.23, 0.83) *	0.47 (0.25, 0.89) *
Education Level		
Less than high school	0.79 (0.61, 1.00)	0.78 (0.60, 1.00)
High school	1.19 (0.94, 1.52)	1.17 (0.93, 1.47)
High school above	0.63 (0.43, 1.00)	0.63 (0.42, 1.00)
Alcohol Consumption		
Never	0.81 (0.66, 0.99) *	0.80 (0.65, 0.99) *
< 1 time/month	2.86 (1.11, 7.35) *	2.53 (1.06, 6.05) *
2~4 times/month	0.95 (0.64, 1.42)	0.97 (0.66, 1.42)
> 4 times/month	0.96 (0.68, 1.34)	0.98 (0.71, 1.37)
Hemoglobin Level		
Quartile 1	0.88 (0.62 ,1.23)	0.87 (0.63 ,1.21)
Quartile 2	0.96 (0.70, 1.31)	0.93 (0.68, 1.26)
Quartile 3	0.74 (0.53, 1.01)	0.74 (0.54, 1.01)
Quartile 4	1.06 (0.77, 1.47)	1.09 (0.80, 1.50)
Serum Ferritin Level		
Quartile 1	0.71 (0.52, 0.97) *	0.70 (0.51, 0.96) *
Quartile 2	0.87 (0.63, 1.19)	0.86 (0.63, 1.19)
Quartile 3	0.91 (0.66, 1.24)	0.90 (0.66, 1.24)
Quartile 4	1.12 (0.82, 1.52)	1.09 (0.80, 1.48)
Serum Folate Level		
Quartile 1	0.85 (0.61, 1.18)	0.82 (0.59, 1.13)
Quartile 2	0.97 (0.71, 1.33)	0.95 (0.70, 1.29)
Quartile 3	1.24 (0.90, 1.71)	1.22 (0.90, 1.68)
Quartile 4	0.62 (0.50, 0.86) *	0.63 (0.46, 0.87) *
Serum Vitamin A Level		
Quartile 1	0.87 (0.62, 1.23)	0.90 (0.65, 1.26)
Quartile 2	1.05 (0.77, 1.43)	1.00 (0.74 ,1.36)
Quartile 3	0.74 (0.54, 1.02)	0.78 (0.57, 1.06)
Quartile 4	0.90 (0.64, 1.26)	0.90 (0.65, 1.27)

Table S6. (Continued).

Effect modifier	Complete data	Multiple imputation
Serum Vitamin E Level		
Quartile 1	0.92 (0.66, 1.30)	0.89 (0.64, 1.24)
Quartile 2	0.80 (0.56, 1.11)	0.86 (0.62, 1.81)
Quartile 3	0.92 (0.68, 1.26)	0.88 (0.64, 1.19)
Quartile 4	0.98 (0.72, 1.34)	0.98 (0.72, 1.33)
Serum β-Carotene Level		
Quartile 1	1.00 (0.71, 1.44)	0.98 (0.69, 1.41)
Quartile 2	1.00 (0.74, 1.34)	1.00 (0.76, 1.33)
Quartile 3	0.88 (0.64, 1.22)	0.85 (0.62, 1.16)
Quartile 4	0.75 (0.55, 1.04)	0.76 (0.55, 1.05)

Note: Factors were not adjusted when they were used as effect modifiers.

[1] Complete data did not include missing data.

[†] The model failed because of the small sample size of other covariates in this stratum.

* Statistically significant associations ($p < 0.05$).