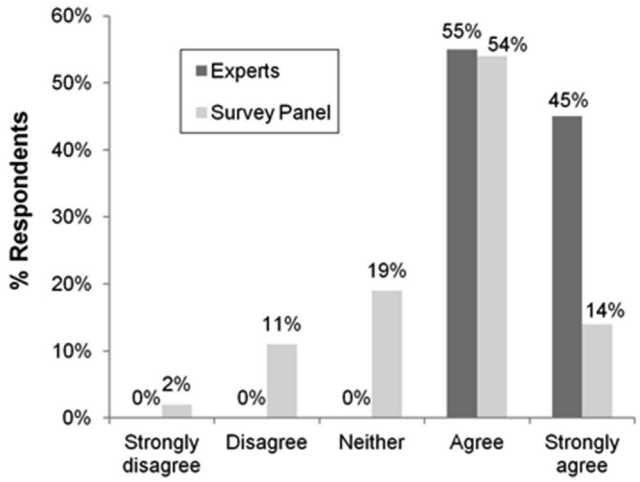
Potential Issues With Serology and the Problem of False Positive Tests

With any test, interpretation depends on the sensitivity and specificity as well as pretest probability. For example, when used in screening in asymptomatic individuals, or in low *H pylori* prevalence areas, many positive tests will be false positive results. Consider a population of 1000 asymptomatic US individuals in which the prevalence of *H pylori* is 15%. Available tests vary in sensitivity and specificity ranging from a high of approximately 95% for the urea breath test or a second-generation monoclonal-based stool antigen immunoassay to a low of 70% or less for some in-house serologic tests. Supplementary Table 3 shows the results of testing depending on the sensitivity and specificity of the available tests. The table shows the positive and negative likelihood ratios, the number of positive and negative test results, and the proportion that are true positive and true negative. Even with the most sensitive and specific tests (eg, 95%) only 77% of positive tests are correct and this proportion rapidly falls as the specificity and sensitivity decrease. For example, a serologic test that is 75% sensitive and specific would produce 324 positive results, of which only 150 were true positives (http://araw.mede.uic.edu/cgi-bin/testcalc.pl). Supplementary Figure 7 shows the marked effect of disease prevalence on the positive and negative predictive values, which is reflected in most negative tests being correct in low disease prevalence situations. If the true prevalence rate were only 4%, only 20% of positives would be true positives. In populations with a high pretest probability such as a population of duodenal ulcer patients, the results would be reversed and we would need to question whether negative tests were correct. For example, in 1000 patients with a pretest probability of *H pylori* duodenal ulcer of 85%, we would have 744 positive tests, of which 97% would be true positives, and 256 negative tests, of which 50% would be false negatives (95% confidence interval, 46%–54%). The marked effect of disease prevalence on positive and negative predictive values results in the recommendation to confirm all positive tests in conditions of low pretest probability and all negative tests when the pretest probability is high. In the study mentioned above only 4.2% of the serologic tests were positive.[7] Confirmation requires that the second test differ in mechanism from the original one,[8] such that serology would be confirmed with a urea breath test, stool antigen immunoassay, or histology. This problem is not restricted to serology and extends to all diagnostic testing.

To continue with the example above where the disease prevalence is 15%, treating all the positives would result in treatment of an additional group of false positive cases which would increase the number treated by 33% with the 95% specific or sensitive test, 50% with the 85% specific or sensitive test and 65% specific or sensitive with the 75% test. In theory, to prevent this antibiotic misuse, all positive tests should be confirmed before treatment except in populations with a high pretest probability. Retesting of any positive result has been suggested.[8] However, this provides a poor solution. For example, with a 75% sensitive or specific test and 15% *H pylori* prevalence among 1000 patients would yield 334 positive tests and only 150 *H pylori*–infected (Supplementary Table 1). The positive tests would include only 112 of the 150 with *H pylori* infections, as 38 (25%) would have been scored as false negative and never retested. Although confirmation retesting the group with positive tests (pretest probability 34%) with a 95% sensitive or specific test would identify the majority of the 112 with *H pylori* infection, the overall treatment rate for the 150 infected would be $<$75%. Overall, even tests with high sensitivity and specificity perform relatively poorly in low pretest probability conditions, while tests with a sensitivity and specificity below 90% should be avoided.

In the United States, large commercial testing laboratories may utilize "in-house" developed tests that are not Food and Drug Administration (FDA) approved.[7] Therefore, in the United States one should always demand that diagnostic laboratories use only FDA-approved tests. One clue to identifying unapproved tests is a report showing results for *H pylori* IgG, IgA, and IgM. IgA and IgM *H pylori* diagnostic tests are rarely FDA approved; generally only IgG tests provide reliable results.

A.



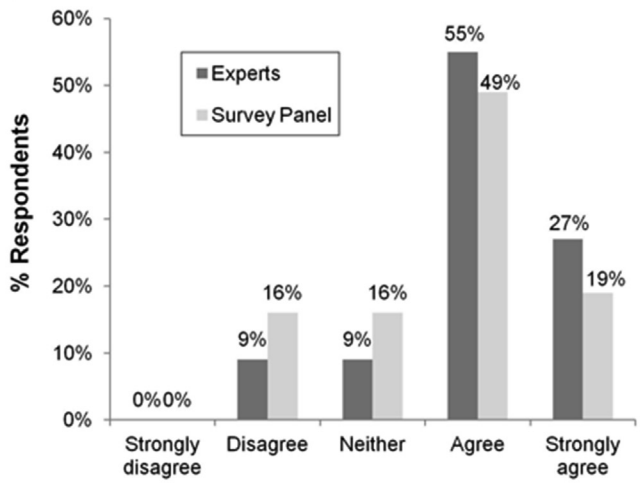
B.

Reasons "Disagree"
- Not indicated
- No need
- Yet to be determined on evidence of medicine
- Low yield
- No clear evidence
- Low risk
- Treatment does not improve ITP
- No significant association

Reasons "Neither Agree nor Disagree"
- Not needed
- Don't think there is need to test
- No strong opinion
- Not enough data for testing
- Not sure if related.
- Unaware of this recommendation
- No data
- Do not have enough info to answer this question
- Is there any good data to support testing these patients?

Supplementary Figure 1. Statement: We recommend that patients with idiopathic thrombocytopenia be tested for *H pylori* infection. (*A*) Proportion responding to each category from strongly disagree to strongly agree. (*B*) Reasons given for responses. Agree/strongly agree: experts, 100%; survey respondents, 68%.

A.



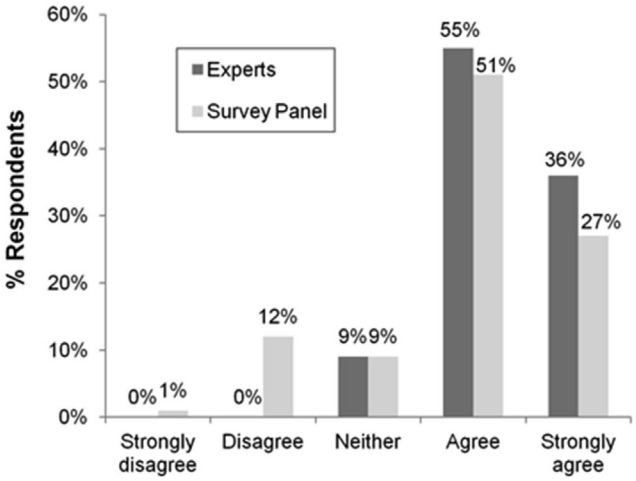
B.

Reasons "Disagree"
- Not indicated
- No need
- No clinical indication
- Not required
- Many patients take PPIs for GERD which have no direct relation to HP
- No clear evidence
- No data
- Unnecessary testing

Reasons "Neither Agree nor Disagree"
- Don't think there is a need to test
- Not sure
- No documentation that this is correct
- Not aware of this as an accepted recommendation
- Is there any good data to support testing these patients?

Supplementary Figure 2. Statement: We suggest that patients receiving long-term PPIs (>1 month) may be tested for *H pylori* infection. (*A*) Proportion responding to each category from strongly disagree to strongly agree. (*B*) Reasons given for responses. Agree/strongly agree: experts, 82%; survey respondents, 68%.

A.



B.

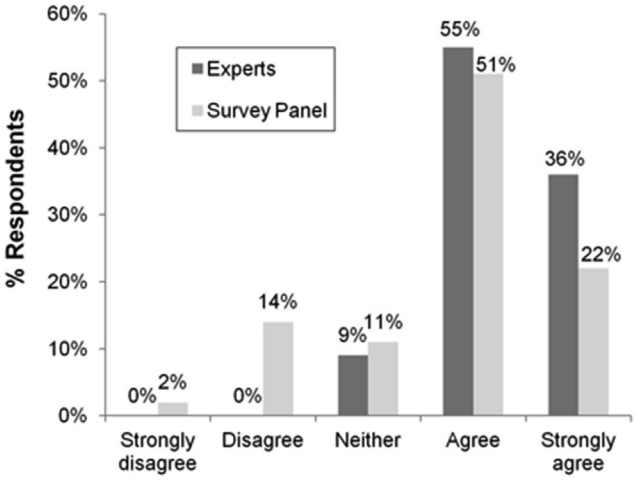
Reasons "Strongly Disagree"
- Not a guideline yet

Reasons "Strongly Disagree"
- Not indicated
- Not proven
- No clinical data to support this
- Yield low
- Only if they are symptomatic
- No data
- Unnecessary testing

Reasons "Neither Agree nor Disagree"
- Pending symptoms
- No strong indication to do so
- Not economically efficient
- Is there any good data to support testing these patients?
- No clear guidelines

Supplementary Figure 3. Statement: We recommend that family members residing in the same household of patients with proven active *H pylori* infections undergo *H pylori* testing. (*A*) Proportion responding to each category from strongly disagree to strongly agree. (*B*) Reasons given for responses. Agree/strongly agree: experts, 91%; survey respondents, 78%.

A.



B.

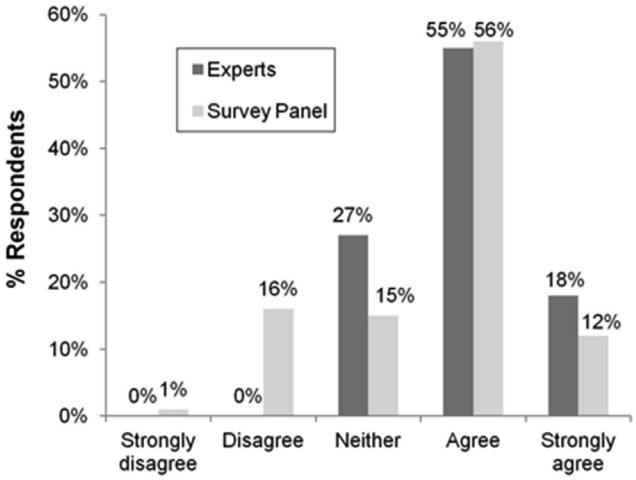
Reasons "Strongly Disagree"
- Not economically efficient
- Not a guideline yet

Reasons "Strongly Disagree"
- Not necessarily needed
- Lacking evidence
- Only if they are symptomatic
- Not necessary
- Unnecessary testing

Reasons "Neither Agree nor Disagree"
- Pending symptoms
- No clinical data to support this
- Unaware of this recommendation
- Sometimes I will and sometimes I won't and since I am probably their parents doctor, If the parents are *H. pylori* negative I won't test the children unless I feel they need it
- Is there any good data to support testing these patients?

Supplementary Figure 4. Statement: We recommend that individuals with family history of peptic ulcer disease be tested for *H pylori* infection. (*A*) Proportion responding to each category from strongly disagree to strongly agree. (*B*) Reasons given for responses. Agree/strongly agree: experts, 91%; survey respondents, 73%.

A.



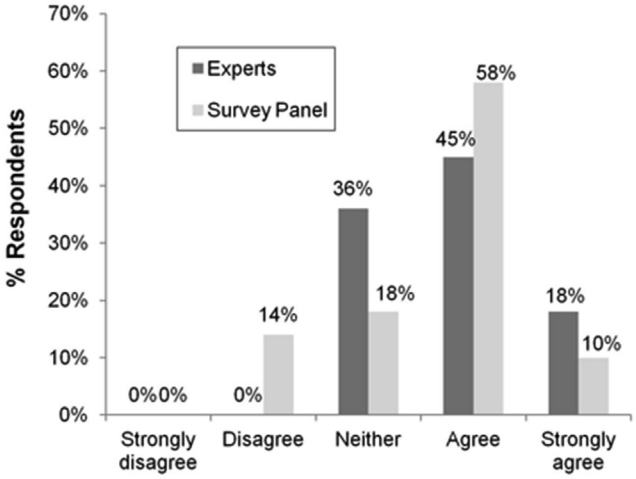
B.

Reasons "Disagree"
- Not needed/ Not necessary
- Many patients on NSAIDs have no risk factors for HP
- Not unless they have symptoms
- No clear evidence
- Low risk

Reasons "Neither Agree nor Disagree"
- Use of NSAIDs is so common that everyone would be getting an EGD
- Don't think there is need to test
- No formal recommendation has been made
- Unaware of this recommendation
- None
- No data
- Is there any good data to support testing these patients?

Supplementary Figure 5. Statement: We suggest that patients taking daily NSAIDs for more than short periods (eg, $>$ 1 month) be tested for *H pylori* infection. (*A*) Proportion responding to each category from strongly disagree to strongly agree. (*B*) Reasons given for responses. Agree/strongly agree: experts, 73%; survey respondents, 68%.

A.

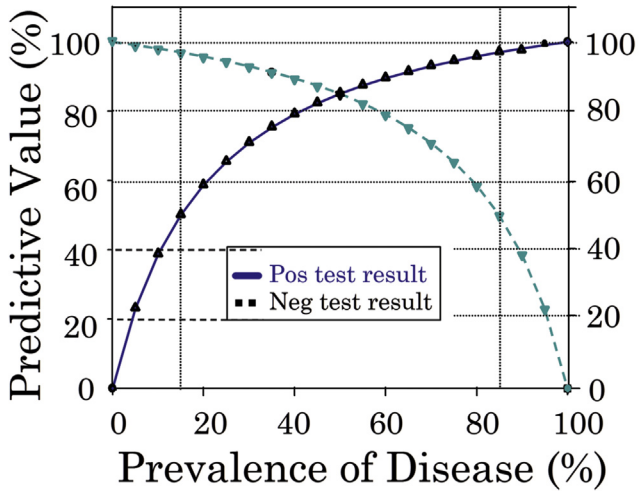


B.

Reasons "Disagree"
- Not indicated
- Not aware of this
- Insufficient data to support this
- If I have trouble titrating a patient on these meds, then I will do it
- No data behind this
- False statement

Reasons "Neither Agree nor Disagree"
- Not needed
- No formal recommendation has been made
- No documentation to this effect
- Unaware of this recommendation
- Don't know the data
- I do not have enough info to answer this question
- Is there good data to support testing these patients?

Supplementary Figure 6. Statement: We suggest that *H pylori* testing be considered in patients treated with medications whose absorption is known to be impacted by infection (eg, l-dopa, thyroxin). (*A*) Proportion responding to each category from strongly disagree to strongly agree. (*B*) Reasons given for responses. Agree/strongly agree: experts, 63%; survey respondents, 68%.



Supplementary Figure 7. The effect of disease prevalence (pretest probability) on the positive and negative predictive values. Results are shown for a test with an 85% sensitivity and specificity and population prevalences of 15% and 85%.

Supplementary Table 1. Criteria for Participation in the Online Survey

- U.S.-based board-certified gastroenterologist
- Years in practice: 5–40 y
- ≥50% of professional time spent in direct patient care
- Provider diagnoses *Helicobacter pylori* in practice
- Provider tests ≥8 patients for *H pylori* per month
- Provider treats ≥4 patients for *H pylori* per month
- Gender: targeted approximately 85% men to match demographics of US gastroenterologist population
- Geography: recruited providers to give equal representation across 4 census regions of United States (Northeast, Midwest, South, and West)
- Excluded anyone who participated in the meeting or who was even invited to the meeting held in Houston, October 27

Supplementary Table 2. Characteristics of the Survey Panel for External Validation

Demographics	
Total number survey respondents	100
Male	86%
Mean age	48 y
Race	
White	52%
Black or African American	1%
American Indian or Alaskan Native	0
Asian	39%
Native Hawaiian or Other Pacific Islander	0
Other	10%
Ethnicity	
Hispanic	4%
Non-Hispanic	96%
Clinical experience	
Mean time in specialty	15.7 y
Time in direct patient care	94.6%
Setting of patient care	
Inpatient	24.4%
Outpatient – solo practice	14.0%
Outpatient – group practice	61.3%
Mean number of patients tested for *Helicobacter pylori* per month	
For initial diagnosis	60
For confirmation of eradication	31
Mean number of patients treated for *H pylori* per month	32

Supplemental Table 3. Tests of *H pylori* Testing in a Population With 1000 Subjects With Tests That Differ in Sensitivity and Specificity

Prevalence (%)	Specificity and sensitivity (%)	Likelihood		Positive	True positive (95% CI) (%)	Negative	True negative (%)
		Positive	Negative				
15	95	19.00	0.05	184	77 (71–82)	816	99
15	85	5.67	0.018	256	50 (46–54)	744	97
15	75	3.00	0.33	324	35 (31–38)	676	94
4	75	3.00	0.33	270	11 (9–13)	730	99
85	75	3.00	0.33	678	94 (93–96)	324	35