# *Helicobacter pylori* Eradication and Gastric Preneoplastic Conditions: A Randomized, Double-Blind, Placebo-Controlled Trial

**Catherine Ley,[1] Alejandro Mohar,[2,3] Jeannette Guarner,[4] Roberto Herrera-Goepfert,[2] Luz Sanchez Figueroa,[5,6] David Halperin,[5] Iain Johnstone,[6] and Julie Parsonnet[1,7]**

[1]Division of Epidemiology, Department of Health Research and Policy, [6]Department of Statistics, and [7]Division of Infectious Diseases and Geographic Medicine, Department of Medicine, Stanford University, Stanford, California; [2]Instituto Nacional de Cancerologia, Mexico City, Mexico; [3]Instituto de Investigaciones Biomedicas, Universidad Nacional Autónoma de México, Mexico City, Mexico; [4]Infectious Diseases Pathology Activity, Centers for Disease Control and Prevention, Atlanta, Georgia; and [5]Division Poblacion y Salud, El Colegio de la Frontera Sur, San Cristobal de las Casas, Chiapas, Mexico

## Abstract

***Helicobacter pylori* causes gastric adenocarcinoma; whether treatment of *H. pylori* infection prevents this cancer remains unknown. In a randomized, double-blind, placebo-controlled trial of *H. pylori* eradication, we determined whether treatment for *H. pylori* decreases gastric cancer risk, using preneoplastic conditions as surrogate markers. A total of 248 healthy volunteers (age >40 years) randomly received *H. pylori* treatment (omeprazole, amoxicillin, clarythromycin; $n = 122$) or matched placebo ($n = 126$) for 1 week. Endoscopy was performed at baseline and at 6 weeks and 1 year. Seven biopsies from each endoscopy were reviewed by two pathologists using the revised Sydney classification. Outcome measures were both a consensus "worst biopsy" diagnosis and a weighted index score that incorporated degrees of severity of preneoplasia from all biopsies. We compared change in these outcomes over time between the two treatment groups. *H. pylori* cure rates for compliant subjects in the treatment arm were 79.2% and 75.7% at 6 weeks and 1 year, respectively. No statistically significant change in the worst biopsy diagnosis was observed from 6 weeks to 1 year between placebo and treated subjects (for improvement/worsening, placebo, 19.4%/10.5%; treatment, 22.5%/8.3%; $P = 0.74$). Change in index score was favorably greater in treatment compared with placebo subjects (intention-to-treat analysis, $P = 0.03$); this finding was particularly evident in the antrum. *H. pylori* eradication gave more favorable gastric histopathologies over 1 year than no treatment. Such incomplete regression suggests but does not prove that eradication of *H. pylori* decreases cancer risk.**

Received 5/31/03; revised 9/5/03; accepted 9/22/03.
**Grant support:** NIH Grant CA67488.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
**Notes:** D. Halperin is deceased; None of the authors has a conflict of interest with this study or its results.
**Requests for reprints:** J. Parsonnet, MD, Grant Building, Room S156, 300 Pasteur Drive, Stanford University School of Medicine, Stanford, CA 94305-5405. Phone: (650) 725-4561; Fax: (650) 498-7011; E-mail: parsonnt@stanford.edu.

## Introduction

Gastric adenocarcinoma remains among the leading causes of death worldwide (1). Intestinal-type gastric adenocarcinoma, the most common type, is preceded by well-defined conditions, starting with superficial gastritis, followed by chronic atrophic gastritis (AG), intestinal metaplasia (IM), and dysplasia (2, 3). Infection with the bacterium *Helicobacter pylori* fosters the development of gastric adenocarcinoma, with organisms containing the Cag pathogenicity island engendering highest risk (4). By contrast, diets rich in fruits and vegetables and low in nitrates and salts are considered protective (5). Randomized clinical trials directly examining gastric adenocarcinoma prevention through *H. pylori* eradication or antioxidant supplementation are ongoing (6). These studies will take years, great expense, and providence to attain sufficient power. Meanwhile, many people will die from a disease for which a prevention strategy—screening and treatment for *H. pylori*—could be cost-effective (7).

Investigations using intermediate biomarkers as endpoints are cost-effective alternatives to cancer prevention trials, although less definitive (8). Several randomized trials examining *H. pylori* eradication and gastric preneoplasia either have been completed or are under way (9–11). Inherent problems include sampling error in obtaining biopsies and misclassification of histological diagnoses. The optimal follow-up time required to assess clinically significant effects is unknown. Finally, the standard method of analyzing preneoplasia uses a consensus diagnosis of the "worst biopsy" collected; this method considers neither information from multiple biopsies nor disparate opinions between pathologists.

To determine whether *H. pylori* eradication is associated with regression of gastric preneoplastic conditions over 1 year, we performed a randomized, double-blind, placebo-controlled clinical trial in Chiapas, Mexico, a region with high gastric malignancy incidence (12). Subjects were healthy volunteers at high risk of preneoplasia. Endpoints included both change in the consensus worst biopsy diagnosis and change in a new stomach index score. Designed to minimize effects of misclassification and sampling error, this index used data from seven biopsies read by two pathologists.

## Materials and Methods

**Study Participants.** Healthy volunteers were recruited in Chiapas, Mexico. An interviewer-administered questionnaire obtained information regarding demographics, family cancer history, and exclusion criteria (age <40 years; current pregnancy; known alcoholism; allergic reactions to study medication; history of malignancy, gastrectomy or debilitating illness; recent antibiotic use; previous *H. pylori* eradication therapy; specific medication use).

Serological selection criteria designed to maximize identification of patients with preneoplasia included CagA antibody positivity with gastrin levels ≥25 μg/ml. CagA antibodies are markers of *H. pylori* strain virulence, and high gastrin levels are an indicator of increased likelihood of AG (13). Blood samples were tested for antibodies to the CagA protein as reported previously (Ref. 14; antigen courtesy of Oravax, Cambridge, MA) and for fasting serum gastrin concentrations by a competitive double-antibody $^{125}$I-radioimmunoassay (Diagnostic Products Corporation, Los Angeles, CA). All samples were stored at −70°C.

**Protocol.** Eligible subjects were randomized to receive antimicrobial therapy (20 mg of omeprazole twice a day, 1 g of amoxicillin twice a day, 500 g of clarythromycin twice a day) or matched placebo for 1 week (the standard treatment time at trial onset in 1996). Masked active therapy and placebo vials had been randomly assigned a number and delivered in bulk to Chiapas. Once a patient was randomized, a single unmasked investigator at Stanford e-mailed to Chiapas the appropriate vial number to be used. Patient randomization followed the biased-coin method with blocking by age ≥60 years and gender (15).

The study was approved by the Human Subject's Committees at Stanford University (Stanford, CA), the Colegio de la Frontera Sur (Chiapas, Mexico), the Instituto Nacional de Cancerologia (Mexico City, Mexico); and the Centers for Disease Control and Prevention (Atlanta, GA). Written informed consent was obtained at the time of initial screening and again before randomization. A data safety and monitoring board at Stanford University regularly reviewed recruitment and progress of the study.

**Assessment.** Treatment compliance (>90% of pills consumed) and side effects were examined 2 weeks after trial enrollment.

Upper endoscopy was performed before treatment and at 6 weeks and 1 year after treatment. Seven biopsies (three each from the antrum and body and one from the incisura angularis) were systematically collected from prespecified locations for histological examination using jumbo forceps. Subjects with ulcers or gastric masses were excluded and referred for treatment. Biopsies were embedded in paraffin, cut, and stained with H&E. Immunohistochemistry for *H. pylori* was performed when inflammation without organisms was seen. Histological parameters, including AG, were evaluated using the updated Sydney system (16); IM was defined as the presence of goblet cells.

Each biopsy was read independently by two general surgical pathologists. Summary consensus diagnoses were then derived for each subject based on a joint review of all biopsies by both pathologists. Subjects with severe dysplasia were excluded and referred for treatment. Pathologists were blinded to treatment arm and endoscopy stage (baseline, 6 weeks, 1 year).

Using the consensus diagnoses and assuming an ascending scale of preneoplastic conditions (normal < AG < IM < dysplasia), we defined a worst biopsy diagnosis for each subject at each time point. Because diagnoses of mild AG or dysplasia are prone to misclassification (17, 18), these diagnoses were ignored, *e.g.*, the revised worst biopsy diagnosis of a subject with, *e.g.*, "mild AG and nonatrophic gastritis" was "nonatrophic gastritis," whereas that for a subject with "mild dysplasia and mild IM" was "mild IM."

The worst biopsy diagnosis takes into consideration neither extent, severity, nor coexistence of conditions from multiple biopsies. We therefore constructed an index to incorporate these characteristics from all biopsies. On the basis of a literature review, the index was modified in accordance with opinions from three outside experts in gastric histopathology (see "Acknowledgments") and included weights for (*a*) increasing degree of preneoplasia (with dysplasia worse than IM or AG), (*b*) preneoplastic condition severity, (*c*) number of affected biopsies, (*d*) missing information (see "Appendix" for model details). The index score was calculated for each subject at each time point, using all biopsies.

**Statistical Analysis.** Assuming a background 10% regression rate, 240 patients were needed to identify a 3-fold difference in response rates between treatment arms with an $\alpha$ of 0.05 and 97% power. Given an expected 10% dropout rate at each study phase, the trial was designed to enroll at least 300 patients (EpiInfo 6.0).

Changes in worst biopsy diagnosis (defined as worsening, no change, or improvement) and in index scores were examined for each subject for each time interval (baseline to 6 weeks, 6 weeks to 1 year, baseline to 1 year). The index score was examined as a mean across pathologists and for each pathologist independently. We used the Spearman coefficient to assess correlations in score within and between pathologists and to assess correlation in change in score over time between pathologists. We used the Wilcoxon rank-sum test to assess associations between worst biopsy diagnosis and index score.

Because a change in index score before and after treatment could be confounded by marked concurrent changes in inflammation attributable to *H. pylori* eradication in the treatment arm (19), we regarded changes in index score over the intervals baseline to 6 weeks and baseline to 1 year as potentially biased. *A priori*, we therefore chose the interval from 6 weeks to 1 year to be the primary interval for comparing change in the two study arms. Changes observed from 6 weeks to 1 year in the placebo group were considered as the natural history of gastric histopathology (time effect). Changes in index score from 6 weeks to 1 year in the treatment group were thus considered to reflect time effect plus treatment effect.

For the intention-to-treat analysis, we compared placebo and treatment groups based on initial randomization. For the per-eradication-protocol analysis, we compared placebo subjects who remained *H. pylori* positive over time with treatment subjects who were cured of *H. pylori* (negative at 6 weeks and 1 year). Subanalyses evaluated changes in the antrum and body separately. All statistical tests were conducted with SAS (version 8.0) software and were two-tailed.

**Role of the Funding Source.** Abbott Laboratories provided all antimicrobial therapy for the clinical trial free of charge; the company was not involved in study design, data management, data analysis, or interpretation of results.

## Results

**Recruitment, Retention, and *H. pylori* Eradication.** A total of 1344 subjects were screened for enrollment. Of the 1178 (87.7%) subjects who gave blood, 468 (39.7%) were CagA-positive with gastrin levels ≥25 μg/ml; 20 of these subjects (4.2%) declined to participate further. The remaining eligible subjects were sequentially invited to participate until a minimum calculated sample size of 300 was achieved. A total of 316 (67.5%) were thus randomized to placebo or treatment arms [155 (49.0%) and 161 (51.0%) subjects, respectively]. The remaining 132 (28.2%) eligible subjects were not invited to participate. Enrollment and follow-up occurred over the period August 1996 to March 1999.

Demographic characteristics and medical histories were similar between enrolled subjects and subjects who did not

meet serological recruitment criteria, declined to participate, or were not randomized (data not shown). Demographic characteristics were also similar between placebo and treatment subjects (Table 1).

No serious adverse effects of treatment were reported. Minor side effects were common (40.6% overall), but only dysgeusia was more prevalent among treatment compared with placebo subjects (78% *versus* 30%; $P \leq 0.001$). Compliance with treatment was very high (93% overall). Among subjects who completed the study, treatment and placebo subjects were equally compliant (96.7% *versus* 96.0%; $P = 1.00$).

Among enrolled subjects, 68 (21.5%) withdrew from the study, primarily because of dislike of endoscopy (58% of withdrawals). One subject with an early gastric carcinoma was withdrawn by investigators.

*H. pylori* cure rates for compliant subjects in the treatment arm were 79.2% and 75.7% at 6 weeks and 1 year, respectively; among placebo subjects, cure rates were 2.9% and 1.9%, respectively ($P < 0.001$).

**Baseline Histopathology, Stomach Index, and Assessment of Time Effect.** Other than having a greater prevalence of nonatrophic gastritis, treatment subjects were similar to placebo subjects with respect to consensus histological diagnoses at baseline (Table 2). Median index scores at baseline were low. The association between index score and worst biopsy diagnosis was strong, with higher scores associated with a higher worst biopsy diagnosis at all time points ($P < 0.001$).

*Table 1* Frequency of demographic, serological, and medical characteristics in enrolled subjects who completed the study (n = 248)[a]

No serious adverse effects of treatment were reported. Minor side effects were common (40.6% overall), but only dysgeusia was more prevalent among treatment compared with placebo subjects (78% *vs.* 30%; $P \leq 0.001$). Compliance with treatment was very high (93% overall).

| | Placebo ($n = 126$) | Treatment ($n = 122$) | $P$[b] |
|---|---|---|---|
| Age (mean ± SD), years | 52.0 ± 9.76 | 51.0 ± 9.22 | 0.40 |
| Female gender, *n* (%) | 79 (62.7) | 78 (63.9) | 0.84 |
| CagA antibody, *n* (%) | | | |
| Positive | 123 (97.6) | 116 (95.1) | 0.33 |
| Borderline | 3 (2.4) | 6 (4.9) | |
| Diabetic | | | |
| Yes | 10 (7.9) | 6 (4.9) | 0.72 |
| Unknown | 1 (0.8) | 1 (0.8) | |
| Smoking history | | | |
| None | 75 (59.5) | 67 (54.9) | 0.80 |
| Former | 32 (25.4) | 37 (30.3) | |
| Occasional | 14 (11.1) | 12 (9.8) | |
| Regular | 5 (4.0) | 6 (4.9) | |
| History of vitamin use,[c] *n* (%) | | | |
| None | 28 (22.2) | 25 (20.5) | 0.88 |
| Former | 27 (21.4) | 24 (19.7) | |
| Occasional | 67 (53.2) | 67 (54.9) | |
| Regular | 4 (3.2) | 6 (4.9) | |
| History of alcohol use, *n* (%) | | | |
| None | 17 (13.5) | 19 (15.6) | 0.59 |
| Former | 22 (17.5) | 21 (17.2) | |
| Occasional | 83 (65.9) | 74 (60.7) | |
| Regular | 4 (3.2) | 8 (6.6) | |
| Known family history of gastric cancer, *n* (%) | 12 (9.5) | 7 (5.7) | 0.43 |

[a] Sixty-eight enrolled subjects who withdrew were not different from those who completed the study.
[b] *t* Test, $\chi^2$ test, or Fisher's exact test, as appropriate.
[c] Dietary components other than vitamin use and alcohol consumption were not assessed.

*Table 2* Baseline histologic characteristics and stomach index score of enrolled subjects who completed the study by trial arm[a]

| Characteristic | Placebo ($n = 126$) | Treatment ($n = 122$) | $P$ |
|---|---|---|---|
| Gastritis, *n* (%) | | | |
| None | 0 | 1 (0.8) | 0.03 |
| Chronic | 14 (11.1) | 4 (3.3) | |
| Acute | 112 (88.9) | 117 (95.9) | |
| Atrophy, *n* (%) | | | |
| None | 47 (37.3) | 56 (45.9) | 0.22 |
| Mild | 39 (31.0) | 28 (23.0) | |
| Moderate | 36 (28.6) | 30 (24.6) | |
| Severe | 4 (3.2) | 8 (6.6) | |
| IM,[c] *n* (%) | | | |
| None | 60 (48.0) | 61 (50.0) | 0.08 |
| Mild | 18 (14.4) | 26 (21.3) | |
| Moderate | 17 (13.6) | 6 (4.9) | |
| Severe | 30 (24.0) | 29 (23.8) | |
| Dysplasia, *n* (%) | | | |
| None | 106 (84.1) | 112 (91.8) | 0.14 |
| Mild | 18 (14.3) | 8 (6.6) | |
| Moderate | 2 (1.6) | 2 (1.6) | |
| Severe | 0 | 0 | |
| Worst biopsy diagnosis,[d] *n* (%) | | | |
| Normal | 0 | 1 (0.8) | 0.83 |
| Gastritis | 51 (40.5) | 46 (37.7) | |
| Atrophy[e] | 10 (7.9) | 14 (11.5) | |
| IM | 63 (50.0) | 59 (48.4) | |
| Dysplasia[e] | 2 (1.6) | 2 (1.6) | |
| Stomach index score,[f] *n* (%) | | | |
| Median | 0.50 | 0.50 | 0.93 |
| 25%–75% | 0–2.0 | 0–1.9 | |
| Range | 0–13.7 | 0–15.3 | |

[a] Enrolled subjects who withdrew were not different at baseline from those who completed the study.
[b] $\chi^2$ test, Fisher's exact test or median test, as appropriate.
[c] IM, intestinal metaplasia.
[d] Normal < chronic/active gastritis < moderate/severe atrophy < mild/moderate/severe IM < moderate dysplasia.
[e] To prevent misclassification, worst diagnoses of mild atrophy were grouped with gastritis because all of these subjects had underlying inflammation. Worst diagnoses of mild dysplasia were grouped with the next worst condition for each subject.
[f] Average score across both pathologists.

The index scores correlated well between pathologists (correlation at baseline, 6 weeks, and 1 year, respectively; $\rho = 0.71$, 0.81, and 0.88, respectively). Change in score over time, however, had heterogeneous variance between pathologists and was poorly correlated between pathologists. As expected, the highest correlation occurred in the 6 weeks to 1 year interval when inflammation was no longer a confounder ($r = 0.15$, 0.44, and 0.33 for change from baseline to 6 weeks, 6 weeks to 1 year, and baseline to 1 year, respectively). Mean change in index score from baseline to 6 weeks was not statistically different from 0 for the placebo group for either pathologist (pathologist 1, $P = 0.46$; pathologist 2, $P = 0.78$). Mean change from 6 weeks to 1 year was also not statistically different from 0, suggesting no time effect (pathologist 1: $P = 0.98$; pathologist 2, $P = 0.15$).

**Effects of Therapy on Worst Biopsy Diagnosis and Stomach Index.** Worst biopsy diagnoses in placebo and treatment subjects were similar from 6 weeks to 1 year and over all other time intervals (Table 3). In contrast, in the same interval, the mean index score of the two pathologists decreased significantly more in the treatment than in the placebo group in both primary analyses: intention-to-treat analysis (Table 3) and per-eradica-

*Table 3* Intention-to-treat analysis: Change over time in both worst biopsy diagnosis and stomach index score among all enrolled subjects by trial arm ($n$ = 126 placebo, 122 treatment)[a]

| | Baseline to 6 weeks | | | 6 weeks to 1 year | | | Baseline to 1 year | | |
|---|---|---|---|---|---|---|---|---|---|
| | Placebo | Treatment | $P^b$ | Placebo | Treatment | $P$ | Placebo | Treatment | $P$ |
| Worst biopsy diagnosis,[c] $n$ (%) | | | | | | | | | |
| Worsening | 17 (13.7) | 14 (11.7) | 0.85 | 13 (10.5) | 10 (8.3) | 0.74 | 11 (8.7) | 11 (9.0) | 0.98 |
| No change | 84 (67.7) | 85 (70.8) | | 87 (70.2) | 83 (69.2) | | 84 (66.7) | 80 (65.6) | |
| Improvement | 23 (18.6) | 21 (17.5) | | 24 (19.4) | 27 (22.5) | | 31 (24.6) | 31 (25.4) | |
| Change in score (mean ± SD) | | | | | | | | | |
| Average[d] | −0.12 ± 2.51 | 0.27 ± 2.47 | 0.22 | −0.14 ± 1.92 | −0.77 ± 2.37 | 0.03 | −0.23 ± 2.25 | −0.5 ± 2.43 | 0.36 |
| Pathologist 1 | −0.13 ± 2.00 | −0.18 ± 2.42 | 0.86 | 0.10 ± 1.76 | −0.50 ± 1.51 | 0.005 | −0.01 ± 2.30 | −0.69 ± 2.32 | 0.02 |
| Pathologist 2 | −0.11 ± 4.30 | 0.73 ± 3.96 | 0.12 | −0.38 ± 2.58 | −1.0 ± 3.98 | 0.13 | −0.45 ± 3.46 | −0.31 ± 3.28 | 0.74 |

[a] Four subjects (two placebo, two treatment) were missing their second endoscopy.
[b] $\chi^2$ or $t$ test, as appropriate.
[c] Normal < chronic or active gastritis < moderate or severe atrophy < mild, moderate, or severe intestinal metaplasia < moderate dysplasia.
[d] Mean of both pathologists' difference in score. Variances were unequal for all time intervals.

tion-protocol analysis (Table 4). Scores also decreased for the interval baseline to 1 year, although this difference was not statistically significant for either analysis. For the period baseline to 6 weeks, treatment subjects appeared to have a statistically insignificant worsening of score compared with placebo.

To determine whether observed effects were consistent across pathologists, we examined results for the pathologists separately. For pathologist 1, in the interval from 6 weeks to 1 year, index scores decreased significantly more in the treatment than in the placebo group in all analyses: intention-to-treat analysis (Table 3), per-eradication-protocol analysis (Table 4), and antrum ($P = 0.01$); body ($P = 0.01$). For the baseline to 1 year interval, pathologist 1 identified similarly significant findings both overall and in the antrum ($P = 0.03$). In the per-eradication-protocol analysis, the magnitude of difference between treatment and placebo arms was similar to that observed in the intention-to-treat analysis although no longer statistically significant. Some regression (defined as any decrease in score) occurred in 39% and 43% and progression in 27% and 18% of placebo and treatment subjects, respectively.

Pathologist 2 observed a magnitude of effect similar to the effect observed by pathologist 1 from 6 weeks to 1 year in both the intention-to-treat and the per-eradication-protocol analyses, but these differences were not statistically significant (Tables 3 and 4, respectively). A statistically significant decrease in score was observed in the antrum ($P = 0.01$). Pathologist 2 did not identify any trend in difference between arms for the period from baseline to 1 year but identified a substantial worsening of score only in the treatment arm from baseline to 6 weeks. Overall, for pathologist 2, some regression occurred in 37% and 37% and progression in 29% and 25% of placebo and treatment subjects, respectively.

## Discussion

In Chiapas, a region with high gastric cancer incidence and high *H. pylori* prevalence, *H. pylori* eradication led to global improvement in gastric histopathology after 1 year as measured by a stomach index score. Regression was evident, particularly in the antrum compared with the body of the stomach. Although improvement was statistically more evident for one pathologist than for the second, both observed a similar magnitude of effect. These results suggest that *H. pylori* eradication, even in older adults, may reduce gastric cancer risk.

Although most subjects showed no change in worst biopsy diagnosis, neither regression nor progression of conditions over 1 year were rare over this short time interval. This may be because of intrinsic sampling biases or because of true change. Although our study follow-up was relatively short (48 week on average in our principle analysis), given the rate of turnover of the gastric mucosa, 1 year would represent dozens of growth cycles and provides ample opportunity for regression. Moreover, a study of 3399 Chinese subjects that examined the natural history of precancerous conditions over 4.5 years found

*Table 4* Per-eradication-protocol analysis: Change over time in both worst biopsy diagnosis and stomach index score among enrolled subjects who met eradication protocol by trial arm ($n$ = 104 placebo, 75 treatment)[a]

| | Baseline to 6 weeks | | | 6 weeks to 1 year | | | Baseline to 1 year | | |
|---|---|---|---|---|---|---|---|---|---|
| | Placebo | Treatment | $P^b$ | Placebo | Treatment | $P$ | Placebo | Treatment | $P$ |
| Worst biopsy diagnosis,[c] $n$ (%) | | | | | | | | | |
| Worsening | 15 (14.4) | 9 (12.0) | 0.76 | 12 (11.5) | 6 (8.0) | 0.47 | 11 (10.6) | 9 (12.0) | 0.96 |
| No change | 68 (65.4) | 53 (70.7) | | 75 (72.1) | 52 (69.3) | | 68 (65.4) | 48 (64.0) | |
| Improvement | 21 (20.2) | 13 (17.3) | | 17 (16.4) | 17 (22.7) | | 25 (24.0) | 18 (24.0) | |
| Change in score (mean ± SD) | | | | | | | | | |
| Average[d] | −0.20 ± 2.62 | 0.27 ± 2.37 | 0.22 | −0.07 ± 1.99 | −0.84 ± 2.21 | 0.02 | −0.27 ± 2.42 | −0.57 ± 2.15 | 0.39 |
| Pathologist 1 | −0.18 ± 2.13 | −0.11 ± 2.42 | 0.85 | 0.15 ± 1.88 | −0.52 ± 1.73 | 0.02 | −0.03 ± 2.44 | −0.63 ± 2.25 | 0.10 |
| Pathologist 2 | −0.22 ± 4.49 | 0.66 ± 4.04 | 0.18 | −0.20 ± 2.55 | −1.17 ± 3.42 | 0.06 | −0.50 ± 3.77 | −0.51 ± 3.02 | 0.99 |

[a] Baseline *H. pylori* positive by histology and either (*a*) *H. pylori* positive at 6 weeks and 1 year if in the placebo group or (*b*) *H. pylori* negative at 6 weeks and 1 year if in the treatment group.
[b] $\chi^2$ test or $t$ test, as appropriate.
[c] Normal < chronic or active gastritis < moderate or severe atrophy < mild, moderate, or severe intestinal metaplasia < moderate dysplasia.
[d] Mean of both pathologists' difference in score. Variances were unequal for all time intervals.

similar degrees of change (20); regression of conditions was observed in 28–88% of subjects (depending on baseline worst biopsy diagnosis) and progression or worsening was also common (4–41%, again dependent on baseline diagnosis). Our results support the finding that the stomach's background rate of change is relatively high. In the end, our results are not substantially different in magnitude of effect from those with longer follow-up (10, 11).

Given this high background rate of change, randomization in therapeutic trials that assess preneoplasia prevention is essential. Even with randomization, however, changes attributable to treatment may be difficult to identify, particularly if small. Sampling error in our study was minimized by the systematic collection of seven biopsies, providing a picture of preneoplasia in the entire stomach. Misclassification was examined in two ways. First, we did not include "soft" diagnoses (mild AG and mild dysplasia) in our analyses. These diagnoses are difficult to make, particularly in the context of inflammation, and agreement among pathologists is often poor (21–23). By omitting these diagnoses, we are more confident that observed differences reflected true changes in preneoplasia rather than merely reduced inflammation. Second, we examined index scores separately for each pathologist. The correlation of scores between pathologists was high, suggesting that global interpretations of pathology were similar even when individual biopsy slide diagnoses were not. Moreover, in a prospective population-based cohort study of gastric cancer conducted in a different study population with different pathologists, increasing score was found to be associated with development of cancer; each one point increment score indicated an ~20% increase in cancer risk over 5 years.[8]

Both pathologists showed an important effect of therapy in the antrum. This localized effect can be explained by the higher frequency of advanced conditions associated with *H. pylori* in this anatomical region (24). Why one pathologist identified a treatment effect more consistently than the second in other regions requires speculation. We believe the observed differences may reflect focused experience by pathologist 1 with gastric conditions. High levels of specialization appear needed to identify modest changes in gastric histopathology.

Other randomized placebo-controlled trials of *H. pylori* eradication and precancerous conditions have used the worst biopsy diagnosis to assess change, including studies involving 631 subjects in Colombia (10) and 587 subjects in China (11). The Colombian study examined the effects of *H. pylori* eradication and dietary supplements. No changes in precancerous conditions were identified for most subjects (54%). Moreover, progression rates were similar across all arms (19–33%). Regression rates, however, were higher in all treatment arms compared with placebo (7% for placebo *versus* 19–29% for treatment arms), and the placebo regression rate was low compared with other studies (20). Improvements were reported only after 6 years of follow-up and occurred in few subjects. In the Chinese trial, a minority of subjects had changes in their worst biopsy score of either AG or IM after 1 year. Eradication of *H. pylori*, however, produced less progression in degree of AG and small decreases in degree of antral IM. Taken together with our study, these randomized trials from three different continents present a consistent picture of improvement in gastric preneoplasia with *H. pylori* eradication, but only in a small subset of subjects.

[8] A. Llosa, Stanford University, and M. Gail, National Cancer Institute, NIH, unpublished data.

Studies of intermediate biomarkers such as these provide circumstantial evidence that *H. pylori* eradication diminishes gastric cancer risk. However, modest improvements in histopathology without total regression cannot be imputed to have clinical significance (8). It remains plausible that only clinically innocuous lesions were reversed with *H. pylori* eradication and that malignancy would occur unchecked. On the basis of these results, screening and treatment programs to eradicate *H. pylori* cannot be mandated, and definite proof that *H. pylori* eradication prevents progression to gastric cancer will require the completion of large randomized trials using cancer as the outcome. Meanwhile, examination of the molecular underpinnings of changes in preneoplastic conditions, *e.g.*, decreases in deleterious mutations, oncogene expression, or microsatellite instability, might bolster the argument that improvements are clinically significant.

Until the time (if ever) that randomized clinical trials of cancer prevention are completed, physicians need to make clinical decisions for their *H. pylori*-infected patients. The authors of several studies have argued that *H. pylori* screening and treatment may be a cost-effective cancer prevention strategy in middle-aged adults, even if treatment prevented only 20–30% of *H. pylori*-associated cancers (7, 25). Eradication therapy is highly effective, well tolerated, and inexpensive. The European Consensus Guidelines, unlike those from the NIH, recommend *H. pylori* eradication therapy for a wide range of symptomatic and asymptomatic infected subjects, including those with preneoplasia (26, 27). Chronic inflammation in various organ systems can lead to malignancy. Without question, in our study and countless others, *H. pylori* eradication mitigates gastric inflammation. Intuitively, this finding alone would suggest that benefit accrues from eradication therapy. This information, combined with studies on preneoplastic regression and evidence of decreased oxidative stress and cell proliferation with *H. pylori* eradication (28), suggest that *H. pylori* eradication could be reasonably applied more widely in adults pending definitive randomized trials of caner prevention.

## Acknowledgments

We thank Drs. Robert Genta, Michael Dixon, and Ernst Kuipers for expert opinions in model design, and Dr. Philip Lavori for providing a statistical perspective. We gratefully thank Shufang Yang at Stanford for technical assistance, and Raul Belmonte, Cecilia Limón, Juan Antonio Moguel, and Rosario Moreno in Comitan for assistance with data collection. We are particularly grateful to El Centro de Investigaciones en Salud de Comitan and El Colegio de la Frontera Sur in Chiapas, Mexico, for the use of their facilities.

## Appendix: Model Construction

The results from each of the seven biopsies from the *i*th patient produced a vector of counts $nn_i = (n_{ATROi}, n_{IMi}, n_{DYSPi})$ of the number of occurrences of atrophy (ATRO), IM, and, dysplasia (DYSP). To create a summary score for statistical analysis, the result of each batch of seven biopsies on the *i*th patient was coded by the index:

$$\begin{aligned} Y_i = \alpha_{ATRO} * [&\beta_{\text{moderate ATRO}} * (n_{\text{moderate ATROi}})^b \\ &+ \beta_{\text{severe ATRO}} * (n_{\text{severe ATROi}})^b] \\ &+ \alpha_{IM} * [\beta_{\text{mild IM}} * (n_{\text{mild IMi}})^b \\ &+ \beta_{\text{moderate IM}} * (n_{\text{moderate IMi}})^b \\ &+ \beta_{\text{severe IM}} * (n_{\text{severe IMi}})^b] \\ &+ \alpha_{DYSP} * [\beta_{\text{moderate DYSP}} * (n_{\text{moderate DYSPi}})^b] \end{aligned}$$

Thus each state (atrophy, IM, and dysplasia) was assigned a severity weight $\alpha$. Additionally, within each state, degree weights $\beta$ were assigned to each degree (mild, moderate, or severe). Severity and degree weights were assigned based on published data of the risk of gastric cancer given the various conditions and on a

*Table 5* Stomach index scores based on seven biopsies for a series of hypothetical profiles[a]

| Biopsies with preneoplastic findings | Stomach index score |
|---|---|
| Unaffected | 0.00 |
| One mild IM[b] | |
| One moderate atrophy | |
| One moderate IM | 0.50 |
| One severe atrophy | |
| One severe IM | |
| Two mild IM | 1.15 |
| Two moderate atrophy | 1.41 |
| Two moderate IM | 1.55 |
| One moderate atrophy/one mild IM | 1.82 |
| One mild IM/one moderate IM | 1.91 |
| Six mild IM | 2.00 |
| Two severe atrophy | 2.00 |
| Four moderate atrophy | 2.00 |
| Two severe IM | 2.00 |
| Five moderate IM | 2.45 |
| Three severe IM | 2.45 |
| One moderate atrophy/two moderate IM | 2.55 |
| Seven moderate atrophy | 2.65 |
| Four severe atrophy | 2.83 |
| One mild IM/one moderate IM/one severe IM | 3.33 |
| Two severe atrophy/two moderate IM | 3.55 |
| Three moderate atrophy/three moderate IM | 3.63 |
| Seven severe IM | 3.74 |
| Seven severe atrophy | 3.74 |
| Seven severe atrophy/seven severe IM | 7.48 |
| One moderate dysplasia | 10.00 |
| Two moderate dysplasia | 14.14 |
| Seven severe atrophy/seven severe IM/seven moderate dysplasia | 33.94 |

[a] If a single biopsy of seven total was affected with moderate or severe atrophy or IM of any degree, then score $Y_i = 0.5$. Otherwise, $Y_i = [(1.00 * (n_{moderate\ ATROi})^{0.5} + 1.4142 * (n_{severe\ ATROi})^{0.5}] + 2.70 * [(0.3024 * (n_{mild\ IMi})^{0.5} + 0.4057 * (n_{moderate\ IMi})^{0.5} + 0.5238 * (n_{severe\ IMi})^{0.5}] + 10.00 * (1.00 * n_{moderate\ DYSPi})^{0.5}]$, where $n$ is the number of biopsies with each condition.
[b] IM, intestinal metaplasia.

synthesis of opinions from expert pathologists, respectively (see below). The total number of occurrences or biopsies with each state was then weighted by a power $b$ to take into consideration the extent of each state (see below).

On the basis of the limited data that quantify the risk of gastric cancer given preneoplastic conditions (3, 20, 29–32), severity weights for atrophy, IM, and dysplasia were conservatively defined as 1.00, 2.70, and 10.00, respectively. Occurrences of unaffected biopsies and biopsies with gastritis alone were not considered in the model because of the lack of information on their association with gastric cancer. Additionally, *H. pylori* eradication therapy is correlated with a decrease in acute inflammation (19); an index that included inflammation or gastritis alone would automatically generate lower scores for active treatment compared with placebo subjects.

To estimate degree weights $\beta$, five rules were defined based on opinions from experts. Given seven biopsies (*a*) one biopsy with dysplasia was always worse than extensive (*i.e.*, seven biopsies with) IM and extensive atrophy; (*b*) extensive IM was equivalent to extensive atrophy; (*c*) two biopsies with severe IM were equivalent to six with mild IM; (*d*) three biopsies with severe IM were equivalent to five with moderate IM; and (*e*) two biopsies with severe atrophy were equivalent to four with moderate atrophy. After we arbitrarily chose $\beta_{moderate\ ATRO}$ as 1.00, the specific degree weights $\beta_{severe\ ATRO}$, $\beta_{mild\ IM}$, $\beta_{moderate\ IM}$, $\beta_{severe\ IM}$, and $\beta_{moderate\ DYSP}$ were derived as 1.4142, 0.3024, 0.4045, 0.5238, and 1.00, respectively.

The presence of a single biopsy with atrophy or IM of any degree out of seven total biopsies was considered by the experts as a localized abnormality attributable to previous ulceration or erosion. In this case, the assigned score was less than the lowest possible score created from two biopsies (two with mild IM only) and was set arbitrarily at 0.5.

The extent of the condition in the stomach was represented by the number of affected biopsies ($b$). Because the risk of cancer increases in a presumably nonlinear fashion with the number of affected biopsies, we arbitrarily chose $b$ as ½ (*e.g.*, a stomach with two or three affected biopsies was considered to be 1.4 and 1.7 times worse than a stomach with a single affected biopsy, and not 2 and 3 times worse). Table 5 presents typical profiles for a variety of hypothetical subjects with their associated scores.

Not all subjects had results from seven biopsies for each condition. Tissue poorly positioned on a slide may be uninterpretable by the pathologist. To adjust for potential differences in available information, a proportional weight was calculated for each subject. This proportional weight was the total potential weight (all biopsies; this weight is the same for all subjects) over the actual weight contributed from the available biopsies. Thus patients who were missing substantial amounts of information for technical reasons were not penalized in the overall group mean compared with patients with all information.

## References


1. Murray, C. J., and Lopez, A. D. Mortality by cause for eight regions of the world: Global Burden of Disease Study. Lancet, *349:* 1269–1276, 1997.

2. Correa, P., Cuello, C., Duque, E., Burbano, L. C., Garcia, F. T., Bolanos, O., *et al.* Gastric cancer in Colombia. III. Natural history of precursor lesions. J. Natl. Cancer Inst. *57:* 1027–1035, 1976.

3. Correa, P., Haenszel, W., Cuello, C., Zavala, D. E., Fontham, E., Zarama, G., *et al.* Gastric precancerous process in a high risk population: cohort follow-up. Cancer Res., *50:* 4737–4740, 1990.

4. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. *Helicobacter pylori*. Schistosomes, liver flukes and *Helicobacter pylori*: views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, pp. 177–240. Lyon, France: IARC, 1994.

5. Potter, J. D., and Steinmetz, K. Vegetables, fruit and phytoestrogens as preventive agents. IARC Sci. Publ., *139:* 61–90, 1996.

6. Forman, D. Lessons from ongoing intervention studies. *In:* R. H. Hunt and G. N. Tytgat (eds.), *Helicobacter pylori:* Basic Mechanisms to Clinical Cure, 1998, pp. 354–361. Dordrecht: Kluwer Academic Publishers, 1998.

7. Parsonnet, J., Harris, R. A., Hack, H. M., and Owens, D. K. Modelling cost-effectiveness of *Helicobacter pylori* screening to prevent gastric cancer: a mandate for clinical trials. Lancet, *348:* 150–154, 1996.

8. Kelloff, G. J., Sigman, C. C., Johnson, K. M., Boone, C. W., Greenwald, P., Crowell, J. A., *et al.* Perspectives on surrogate end points in the development of drugs that reduce the risk of cancer. Cancer Epidemiol. Biomark. Prev., *9:* 127–137, 2000.

9. Munoz, N., Vivas, J., Buiatti, E., Kato, I., and Oliver, W. Chemoprevention trial on precancerous lesions of the stomach in Venezuela: summary of study design and baseline data. IARC Sci. Publ., *139:* 125–133, 1996.

10. Correa, P., Fontham, E. T., Bravo, J. C., Bravo, L. E., Ruiz, B., Zarama, G., *et al*. Chemoprevention of gastric dysplasia: randomized trial of antioxidant supplements and anti-*Helicobacter pylori* therapy. J. Natl. Cancer Inst. (Bethesda), *92:* 1881–1888, 2000.

11. Sung, J. J., Lin, S. R., Ching, J. Y., Zhou, L. Y., To, K. F., Wang, R. T., *et al*. Atrophy and intestinal metaplasia one year after cure of *H. pylori* infection: a prospective, randomized study. Gastroenterology, *119:* 7–14, 2000.

12. Halperin, D. C., Belgrade, M. E., and Mohar, A. Stomach cancer cluster in Mexico. Lancet, *2:* 1055, 1988.

13. Ley, C., Mohar, A., Guarner, J., Herrera-Goepfert, R., Figueroa, L. S., Halperin, D., *et al*. Screening markers for chronic atrophic gastritis in Chiapas, Mexico. Cancer Epidemiol. Biomark. Prev., *10:* 107–112, 2001.

14. Parsonnet, J., Friedman, G. D., Orentreich, N., and Vogelman, J. H. Risk for gastric cancer in persons with CagA positive and CagA negative *Helicobacter pylori* infection. Gut, *40:* 297–301, 1997.

15. Efron, B. Forcing a sequential experiment to be balanced. Biometrika, *58:* 403–417, 1971.

16. Dixon, M. F., Genta, R. M., Yardley, J. H., and Correa, P. Classification and grading of gastritis. The updated Sydney System. International Workshop on the Histopathology of Gastritis, Houston 1994. Am. J. Surg. Pathol., *20:* 1161–1181, 1996.

17. Guarner, J., Herrera-Goepfert, R., Mohar, A., Sanchez, L., Halperin, D. Ley, C., *et al*. Interobserver variability in application of the revised Sydney classification for gastritis. Hum. Pathol., *30:* 1431–1434, 1999.

18. van Grieken, N. C., Weiss, M. M., Meijer, G. A., Bloemena, E., Lindeman, J., Offerhaus, G. J., *et al*. Rapid quantitative assessment of gastric corpus atrophy in tissue sections. J. Clin. Pathol., *54:* 63–69, 2001.

19. Di Napoli, A., Petrino, R., Boero, M., Bellis, D., and Chiandussi, L. Quantitative assessment of histological changes in chronic gastritis after eradication of *Helicobacter pylori.* J. Clin. Pathol., *45:* 796–798, 1992.

20. You, W. C., Zhang, L., Yang, C. S., Chang, Y. S., Issaq, H., Fox, S. D., *et al*. Evolution of precancerous lesions in a rural Chinese population at high risk of gastric cancer. Int. J. Cancer, *83:* 615–619, 1999.

21. Plummer, M., Buiatti, E., Lopez, G., Peraza, S., Vivas, J., Oliver, W., *et al*. Histological diagnosis of precancerous lesions of the stomach: a reliability study. Int. J. Epidemiol., *26:* 716–720, 1997.

22. Chen, X., van der Hulst, R. W. M., Bruno, M. J., van der Ende, A., Xiao, S., Tytgat, G. N. J., *et al*. Interobserver variation in the histopathological scoring of *Helicobacter*-related gastritis. J. Clin. Pathol., *52:* 612–615, 1999.

23. Tepes, B., Ferlan-Marolt, V., Jutersek, A., Kavcic, B., and Zaletel-Krageli, L. Interobserver agreement in the assessment of gastritis reversibility after *Helicobacter pylori* eradication. Histopathology, *34:* 124–133, 2001.

24. You, W., Blot, W. J., Li, J. Y., Chang, Y., Jin, M., Kneller, R. W., *et al*. Precancerous gastric lesions in a population at high risk of stomach cancer. Cancer Res., *53:* 1317–1321, 1993.

25. Fendrick, A. M., Chernew, M. E., Hirth, R. A., Bloom, B. S., Bandekar, R. R., and Scheiman, J. M. Clinical and economic effects of population-based *Helicobacter pylori* screening to prevent gastric cancer. Arch. Intern. Med., *159:* 142–148, 1999.

26. Malfertheiner, P., Megraud, F., O'Morain, C., Bell, D., Bianchi, P. G., Deltenre, M., *et al*. Current European concepts in the management of *Helicobacter pylori* infection—the Maastricht Consensus Report. The European *Helicobacter Pylori* Study Group (EHPSG). Eur. J. Gastroenterol. Hepatol., *9:* 1–2, 1997.

27. NIH Consensus Development Panel on *Helicobacter pylori* in Peptic Ulcer Disease. *Helicobacter pylori* in peptic ulcer disease. JAMA, *272:* 65–69, 1994.

28. Moss, S. F. The carcinogenic effect of *H. pylori* on the gastric epithelial cell. J. Physiol. Pharmacol., *50:* 847–856, 1999.

29. Rugge, M., Farinati, F., Baffa, R., Sonego, F., Di Mario, F., Leandro, G., *et al*. Gastric epithelial dysplasia in the natural history of gastric cancer: a multicenter prospective follow-up study. Gastroenterology, *107:* 1288–1296, 1994.

30. Rokkas, T., Filipe, M. I., and Sladen, G. E. Detection of an increased incidence of early gastric cancer in patients with intestinal metaplasia type III who are closely followed up. Gut, *32:* 1110–1113, 1991.

31. Wu, M. S., Shun, C. T., Lee, W. C., Chen, C. J., Wang, H. P., Lee, W. J., *et al*. Gastric cancer risk in relation to *Helicobacter pylori* infection and subtypes of intestinal metaplasia. Br. J. Cancer, *78:* 125–128, 1998.

32. Uemura, N., Okamoto, S., Yamamoto, S., Matsumura, N., Yamaguchi, S., Yamakido, M., *et al. Helicobacter pylori* infection and the development of gastric cancer. N. Engl. J. Med., *345:* 784–789, 2001.

# *Helicobacter pylori* Eradication and Gastric Preneoplastic Conditions: A Randomized, Double-Blind, Placebo-Controlled Trial

Catherine Ley, Alejandro Mohar, Jeannette Guarner, et al.

*Cancer Epidemiol Biomarkers Prev* 2004;13:4-10.