American Journal of Gastroenterology
© 2005 by Am. Coll. of Gastroenterology
Published by Blackwell Publishing

ISSN 0002-9270
doi: 10.1111/j.1572-0241.2005.41657.x

## ORIGINAL CONTRIBUTIONS

# Diagnostic and Therapeutic Use of Proton Pump Inhibitors in Non-Cardiac Chest Pain: A Metaanalysis

Filippo Cremonini, M.D., M.Sc., James Wise, M.D., Paul Moayyedi, Ph.D., F.R.C.P., and Nicholas J. Talley, M.D., Ph.D., F.A.C.P.

*Clinical Enteric Neuroscience, Translational and Epidemiological Research Program (C.E.N.T.E.R.) and Mayo Clinic College of Medicine, Rochester, Minnesota; and Gastroenterology Division, McMaster University, Hamilton, Ontario, Canada*

**OBJECTIVES:** To assess (i) the efficacy of short-term proton pump inhibitors (PPIs) in non-cardiac chest pain (NCCP) and (ii) the performance of an empirical short-term treatment with PPI (PPI test) to establish a diagnosis of abnormal acid reflux in NCCP.

**METHODS:** Metaanalysis of English language studies identified by searching MEDLINE (1966–May 2004), EMBASE (1980–May 2004), Cochrane Controlled Trials Register, and abstract books from major gastroenterology meetings (1993–2004). For the metaanalysis of PPI efficacy in NCCP, we selected randomized controlled trials (parallel group and crossover designs) comparing PPI therapy with placebo. For the metaanalysis of PPI test performance, we selected uncontrolled studies comparing the test with a standard reference.

**RESULTS:** Eight studies were included in the PPI efficacy analysis. The pooled risk ratio for continued chest pain after PPI therapy was 0.54 (95% CI 0.41–0.71). The overall number needed to treat was 3 (95% CI 2–4). The pooled sensitivity, specificity, and diagnostic odds ratio for the PPI test *versus* 24-h pH monitoring and endoscopy were 80%, 74%, and 13.83 (95% CI 5.48–34.91), respectively. All studies were small and there was evidence of publication bias or other small study effects.

**CONCLUSION:** PPI therapy reduces symptoms in NCCP and may be useful as a diagnostic test in identifying abnormal esophageal acid reflux.

(Am J Gastroenterol 2005;100:1226–1232)

## INTRODUCTION

Non-cardiac chest pain (NCCP) is a common condition (1) associated with substantial health-care costs and resource utilization (2). Management is difficult because NCCP remains a diagnosis of exclusion that encompasses heterogeneous patient populations. A subset of patients with NCCP will have evidence of abnormal esophageal acid exposure (3, 4) or esophageal dysmotility (5) that can be associated with features of visceral hypersensitivity (6). In other patients presenting with NCCP, musculoskeletal (7) and psychiatric disorders (8) may be important.

As gastroesophageal reflux disease (GERD) has been implicated in NCCP, proton pump inhibitors (PPIs) have been advocated as a simple diagnostic tool to identify patients with an acid-related disorder. A decision analysis based on the outcome from one clinical trial (9) suggested that a cost-effective method of diagnosing NCCP is a therapeutic trial with potent acid suppression (10). However, the magnitude of benefit of PPI therapy in NCCP is uncertain. Furthermore, it is unclear whether PPI therapy is a useful diagnostic tool. We therefore performed a systematic review to determine: (i) whether PPIs are superior to placebo in terms of inducing symptom improvement in NCCP and (ii) whether a therapeutic trial of PPI accurately diagnoses GERD in NCCP patients.

## METHODS

### *Studies Retrieval and Selection, Outcome Measure*

The present metaanalysis follows the QUOROM statement guidelines (11). We performed an electronic search of the databases MEDLINE (1996–May 2004), EMBASE (1980–May 2004), and the Cochrane-controlled trials register (until May 2004), and a manual search of the abstracts from the American Gastroenterology Association meeting proceedings' books (1993–2004). We also conducted a manual search in the reference lists of the papers retrieved. As search terms, we used *non-cardiac, noncardiac, chest pain, undetermined, unexplained*, with any combinations of the above. We did not include review articles, position papers, editorials, commentaries, and book chapters. The criteria considered for study inclusion are given in Table 1. Two investigators (F.C. and J.W.) separately performed the search, selected the studies, and jointly performed data extraction using pre-defined

**Table 1.** Criteria for Inclusion of Studies in the Metaanalysis of Proton Pump Inhibitor (PPI) Treatment in Non-Cardiac Chest Pain (NCCP)

| |
|---|
| Metaanalysis of PPI treatment efficacy in NCCP |
| • Use of proton pump inhibitor as active treatment |
| • Parallel group and crossover trials |
| • Comparison with placebo |
| • Active treatment and placebo arms similar for all demographic features (if parallel-group studies) |
| • Cardiac origin of chest pain excluded by appropriate investigations |
| • Patients blinded to treatment assignments |
| • Intention-to-treat analysis |
| • Available/extractable raw data |
| Metaanalysis of PPI test performance in NCCP |
| • Exclusion of heart disease with appropriate tests |
| • Labeling of patients as NCCP |
| • 24-h pH monitoring (with or without other tests) used as reference standard for diagnosis of gastroesophageal reflux disease |
| • Comparability of data from PPI test to reference standard |

data extraction forms. A senior investigator (P.M.) also performed data extraction and arbitrated in case of any lack of agreement.

### *Statistical Methods*

The outcome measure of response to PPI treatment was defined as a binary variable. Subjects were considered to be "responders" if they reported symptomatic improvement after PPI administration. For the purpose of the primary analyses, we used the criteria for defining a "responder" chosen by the authors of each paper, as these differed slightly across studies. In sensitivity analyses, we considered complete chest pain relief, >50% chest pain improvement, and any chest pain improvement as the definition of response.

The relative risks of not responding to therapy were pooled using a random-effects model (12). The analysis of treatment effect was performed on an intention-to-treat basis on evaluable patients. The number needed to treat (NNT) was calculated using the reciprocal of the pooled absolute risk reduction.

A Funnel scatterplot was drawn to assess the potential for publication bias. We plotted the studies' risk ratios *versus* the square root of the studies' sample sizes to detect asymmetry in the distribution of trials and regressed the individual studies' risk ratios to the respective sample sizes (13). In a funnel plot, larger studies providing a more precise estimate of the true effect of the intervention in question form the spout of a funnel, whereas smaller studies provide less precise estimate, and form the cone of the funnel. A gap in the funnel plot would indicate the potential for publication bias.

For the question regarding performance characteristics of the PPI test, we plotted the sensitivities and specificities obtained for each study included in a summary. Again, we used the author's definition of a positive test (response to PPI) and in sensitivity analyses used complete chest pain relief, >50% chest pain improvement, and any chest pain improvement as the definition of a positive test. We then obtained the diagnostic odds ratio (DOR) for each of the studies included, and calculated the pooled diagnostic odds ratio using a random-effects model. The DOR describes the odds of a positive test result in patients with the abnormal acid reflux compared with the odds of a positive test result in those without this abnormality. A pooled receiver-operator curve was calculated using the pooled DOR and 95% confidence intervals (14).

All calculations were performed with STATA (version 8, Stata corporation, TX) using the metan and metabias commands.

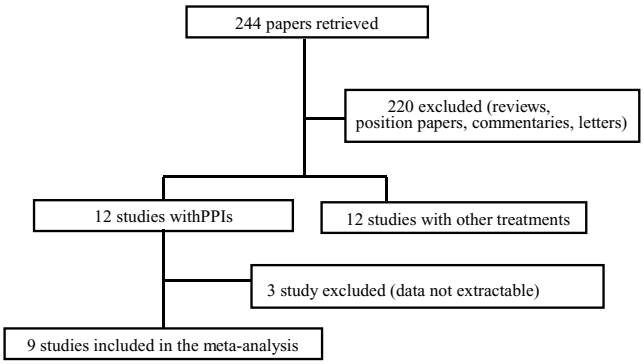


**Figure 1.** Metaanalysis flow.

## RESULTS

### *Study Retrieval and Inclusion*

A flow diagram of this systematic review, with the number of papers retrieved, included, and excluded with the reasons for exclusion, is shown in Figure I. The characteristics of the studies included in the metaanalysis are shown in Table 2.

EFFICACY OF PPI THERAPY IN REDUCING NCCP SYMPTOMS. Seven trials (9, 15–20) were included in the metaanalysis of effect, yielding a total of 232 patients. Two studies were parallel group (15, 20) and the remaining five studies were crossover design. The pooled risk ratio of continued chest pain after PPI treatment was 0.54 (95% CI 0.41–0.71) using the individual studies' definition of response (Fig. 2) giving an NNT of three (95% CI 2 to 4). Sensitivity analyses gave similar results using >50% response and any symptom response as a measure of outcome. Five studies (9, 15, 16, 18, 19) provided data on symptom improvement of >50% with a pooled RR for continued chest pain of 0.60 (95% CI 0.44–0.81). If any improvement in chest pain was used to define the response to PPI, the pooled RR was 0.51 (95% CI 0.33–0.79) in data from five studies (3, 15, 16, 18, 20). There appeared to be less benefit seen with PPI therapy if complete resolution of chest pain was used as an outcome (pooled RR 0.83, 95% CI 0.66–1.05) from the five studies (3, 15, 16, 18, 20) that provided these data (Fig. 3).

The funnel plot showed statistically significant asymmetry with a negative slope (coefficient −1.42; 95% CI = −1.83 to

**Table 2.** Characteristics of Studies Included

| Reference No. | N | PPI, Dose, Duration | Study Design | Proportion with Baseline Characteristics Suggestive of GER (%) | Threshold of Pain Decrease for Defining a "Responder" | Reference Standard (s) |
|---|---|---|---|---|---|---|
| Achem (15) | 36 | Omeprazole 40 mg 8 wk | Placebo-controlled, parallel groups | 100 | "Overall" improvement | None |
| Bautista (16) | 40 | Lansoprazole 90 mg 1 wk | Placebo-controlled, crossover | 45 | >50% | pH-metry, endoscopy |
| Fass (9) | 39 | Omeprazole, 60 mg 1 wk | Placebo-controlled, crossover | 62 | >50% | pH-metry, endoscopy |
| Pandak (18) | 44 | Omeprazole, 80 mg 2 wk | Placebo-controlled, crossover | 53 | >50% | pH-metry, endoscopy |
| Squillace (19) | 17 | Omeprazole 80 mg 1 day | Placebo-controlled, crossover | 76 | >50% | pH-metry |
| Xia (20) | 36 | Lansoprazole 30 mg 4 wk | Placebo-controlled, parallel group | 33 | >50% | pH-metry |
| Fass (17) | 20 | Rabeprazole 40mg 1 wk | Placebo-controlled, crossover | 60 | >50% | pH-metry, endoscopy |
| Dekel (22) | 94 | Various PPIs, 1 wk | Open label | 50 | >50% | pH-metry, endoscopy |
| Chambers (21) | 31 | Omeprazole 40 mg, 6 wk | Open label | 17 | No *a priori* definition | pH-metry |

PPI, proton pump inhibitor; GER, abnormal acid reflux.

−1.01, $p<0.001$). This suggests the potential for publication bias or small study effects (Fig. 4).

ACCURACY OF THE PPI TEST IN DIAGNOSING GORD IN NCCP. For the metaanalysis of diagnostic yield of the PPI test, eight studies were included (9, 16–22) with a total of 321 patients; two studies were nonrandomized (21, 22). The average proportion of participants with esophagitis across the studies was 22% (range 0–43%). Three studies (19–21) used 24-h esophageal pH monitoring alone and five studies (9, 16–18, 22) used 24-h pH monitoring plus endoscopy as a reference standard.

The pooled sensitivity, specificity, and diagnostic odds ratio for the PPI test were 80% (range 0–95%), 74% (range 60–91%), and 13.83 (95% CI 5.48 to 34.91), respectively (Fig. 5). Figure 6 shows the summary ROC curve these data

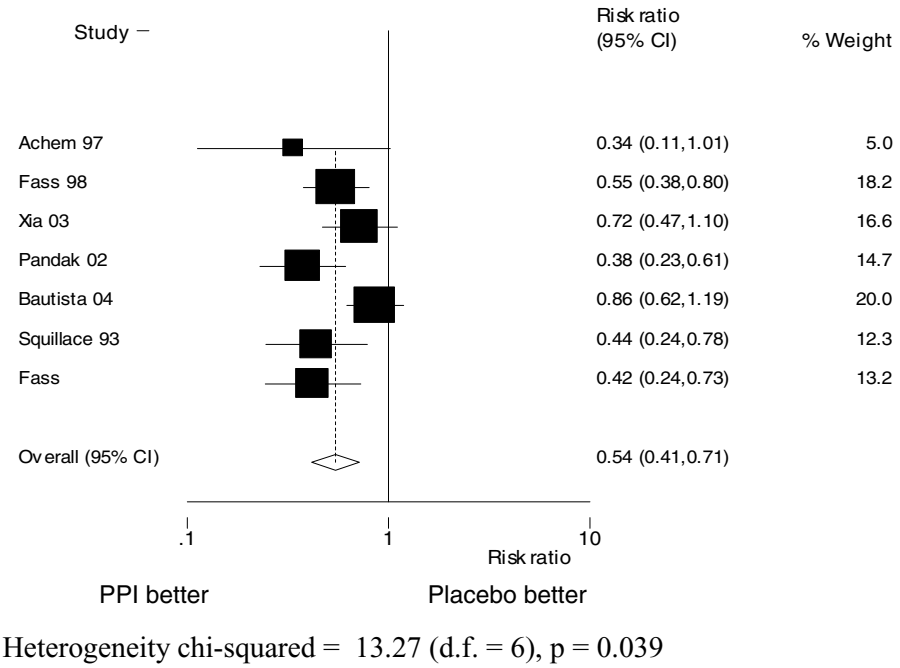


**Figure 2.** Forest plot showing the effect of proton pump inhibitor (PPI) treatment on symptom improvement in NCCP. Estimates of relative risk for continued chest pain after treatment are presented, with their 95% confidence intervals using the authors' definition of response for each study.

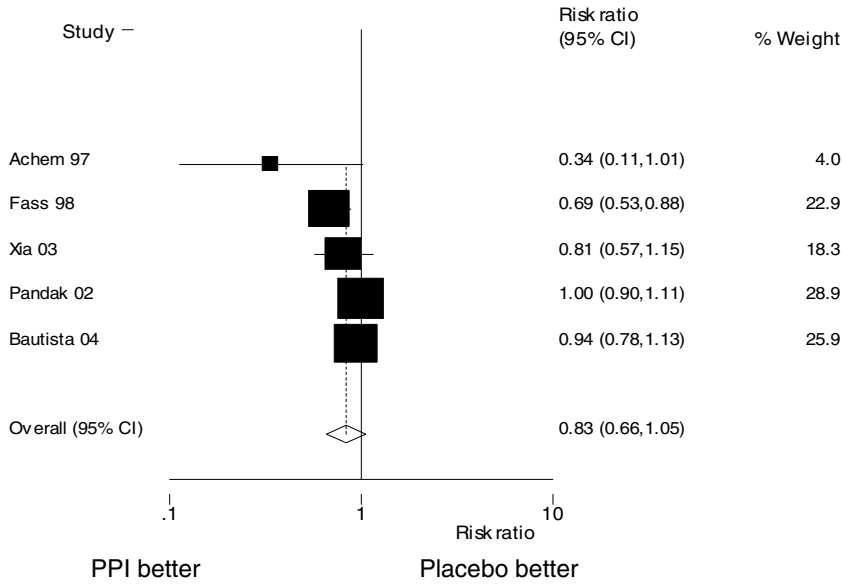


**Figure 3.** Forest plot showing the effect of PPI treatment on complete symptom resolution in non-cardiac chest pain (NCCP). Estimates of relative risk for continued chest pain after treatment are presented, with their 95% confidence intervals using complete symptom resolution as the definition of response for each study.

generated. There was no statistically significant heterogeneity between studies ($\chi^2 = 9.43$ (d.f. $= 7$), $p = 0.223$), and therefore we used these data to calculate a summary ROC curve (Fig. 6). Table 3 shows pooled estimates for sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios using the author's definitions of response. In all studies, there was a high prevalence of esophagitis or heartburn. Table 4 shows the pooled estimates of sensitivity, specificity, and diagnostic odds ratios using different definitions of response to PPIs; any symptom response performed less well as a diagnostic tool than the authors' definitions.

## DISCUSSION

This metaanalysis shows that short-term PPI therapy is superior to placebo in achieving a symptomatic response in patients with NCCP in whom a cardiac etiology has been excluded by appropriate testing, with the number needed to treat being 3. The other main finding is that, in patients with

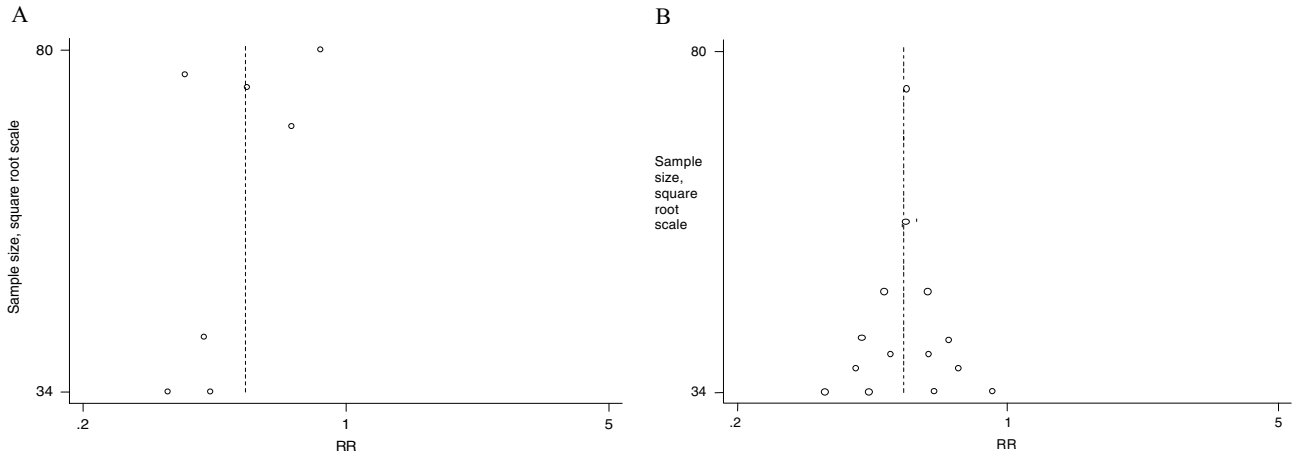


**Figure 4.** Funnel plots for detection of publication bias. (*A*) There is plot asymmetry with a negative slope (coefficient −1.42, 95% CI = 1.83 to −1.01, $p < 0.001$) and significant bias (coefficient = −0.53; 95% CI = −0.91 to −0.16, $p = 0.015$). The plot indicates some of the effect is due to publication bias or small study effects (more likely in this case), with lack of published negative studies with small sample size. (*B*) Sample funnel plot showing no publication bias, with the estimated risks evenly distributed. Larger studies provide more accurate estimates and form the spout of the funnel. Smaller studies provide less accurate estimates and form the cone of the funnel.

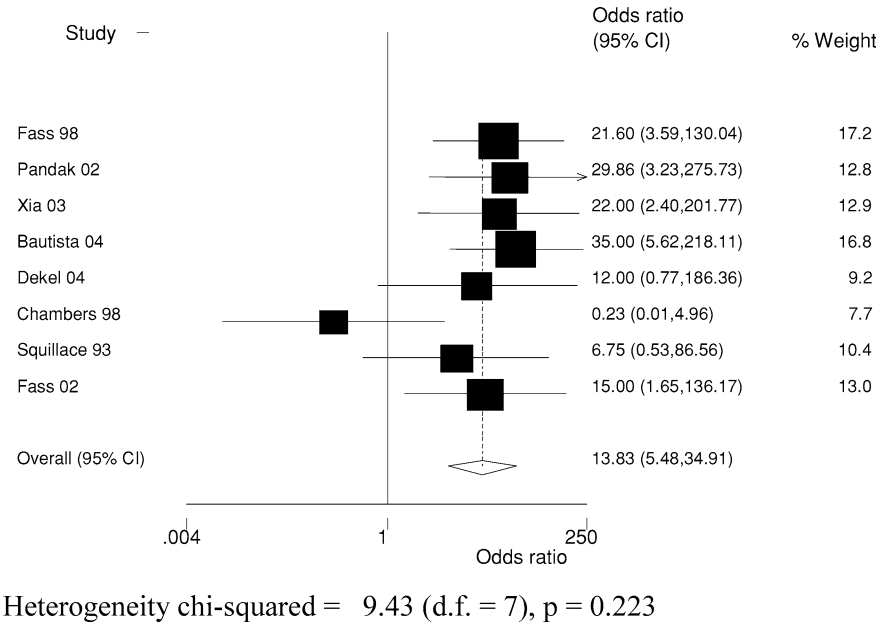


**Figure 5.** Forest plot with individual and pooled diagnostic odds ratios, constructed with the sensitivity and specificity data of the proton pump inhibitor (PPI) test for the presence of abnormal acid reflux in non-cardiac chest pain (NCCP) using the definition of PPI response chosen by the investigators in each study. Heterogeneity refers to the range of effect sizes estimated by the pool of studies. If the test for heterogeneity is significant, there is significant variation between the effect size estimates across studies.

NCCP, the PPI test is a reasonably sensitive and specific test for the presence of abnormal acid reflux as demonstrated by 24-h pH monitoring.

### *Efficacy of PPI Treatment in NCCP*

The use of PPI as initial treatment approach in NCCP is based on the presumption that in a relevant proportion of patients, symptoms are related to the presence of acid reflux.

A cost-effectiveness analysis has shown empirical treatment with PPI is superior to other strategies in NCCP (10). The assumptions of the cost-effectiveness analysis were based on data from a small number of patients. Although the present metaanalysis shows an NNT of three for PPI in NCCP, it should be noted that the success rate of PPI therapy will be largely dependent on the proportion of patients with acid reflux in the population evaluated. Only one study among those included had esophagitis as an exclusion criteria (20), and its results were comparable to the remainder of the studies. Estimates of the prevalence of acid reflux in NCCP (2) from the literature are consistent with that observed in the participants in the therapeutic studies included in the present metaanalysis. These studies were performed in Europe, Asia, and in the United States and enrolled primary care as well as tertiary referral patients. This increases the generalizability of our results, since there were no major differences in the outcomes between different patients' populations. There were, however, only a relatively small number of patients evaluated and there was evidence of publication bias or other small study effects. Hence, we cannot exclude the possibility that the efficacy of PPI therapy may have been overestimated in this analysis.

### *Performance of the PPI Test in NCCP*

Our data suggest that the PPI test could be a useful diagnostic tool for identifying patients with acid reflux-related NCCP,

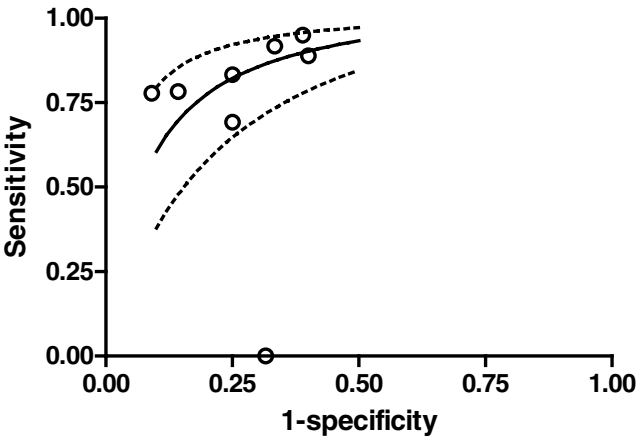


**Figure 6.** Summary receiver operator characteristic (ROC) curve from the pooled diagnostic odds ratio. The overall performance of the test can be determined by the position of the curve. A poor diagnostic test would have a curve whose shape is close to the diagonal, whilst the curve for an ideal tests would be steep and close to the top left corner, where both the sensitivity and the specificity are 1 (14). Each circle represents the results of individual studies. The solid line is an estimate of the pooled ROC curve, with the dotted lines being its 95% confidence intervals.

**Table 3.** PPI Test Performance Summaries According to the Author's Definition and Prevalence of Diagnoses

| Author | Esophagitis (%) | GERD (%)* | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) | LR+ | LR− |
|---|---|---|---|---|---|---|---|---|
| Fass (9) | 43 | 62 | 78 | 86 | 90 | 71 | 5.48 | 0.25 |
| Pandak (18) | 26 | 53 | 95 | 61 | 73 | 92 | 2.44 | 0.08 |
| Xia (20) | 0 | 33 | 92 | 67 | 58 | 94 | 2.75 | 0.13 |
| Bautista (16) | 20 | 45 | 78 | 91 | 88 | 83 | 8.56 | 0.24 |
| Dekel (22) | 35 | NR | 89 | 60 | 80 | 75 | 2.22 | 0.19 |
| Chambers (21) | NR | 17 | 0 | 68 | 0 | 76 | 0 | 1.46 |
| Squillace (19) | 6 | 76 | 69 | 75 | 90 | 43 | 2.77 | 0.41 |
| Fass (17) | NR | 60 | 83 | 75 | 83 | 75 | 3.33 | 0.22 |

*Gastroesophageal reflux disease as defined by a positive 24-h pH study. PPV, positive predictive value; NPV, negative predictive value. LR+, positive likelihood ratio. LR−, negative likelihood ratio. NR, not recorded.

with a pooled sensitivity of 80% and specificity of 74%. Our results should be put in the perspective of the findings of a systematic review of 15 studies that compared the PPI test *versus* 24-h pH monitoring and other reference standards in the diagnosis of reflux disease in patients with upper gastrointestinal symptoms suggestive of GERD (23). This metaanalysis found that the PPI test was of little help for establishing a diagnosis of GERD of with a pooled sensitivity and specificity 78% and 54%, respectively (23). The better performance of the PPI test reported in our metaanalysis might be due to a somewhat higher prevalence of abnormal acid reflux in patients with NCCP. However, as the longest follow-up in the studies included was only 6 wk, we are not able to predict whether the PPI test will help to identify patients more likely to benefit from acid suppression over the long term.

### *Limitations*

Potential limitations do need to be considered. First, the inclusion of crossover trials artificially inflates the studies' sample sizes and this may reduce the estimated 95% confidence interval of the relative risk. Despite most studies having a treatment-washout period that could be deemed of sufficient length, a carryover effect from PPIs to placebo and *vice versa* cannot be excluded. However, such a bias would drive the data toward the null hypothesis. We recommend all future clinical trials of NCCP should consider a parallel group design (24).

Second, the number of studies and patients available was relatively small. Notably, the studies had similar inclusion criteria and measurement of the outcomes. It should also be noted that four of eight studies included were attributable, at least in part, to one research group (9, 16, 17, 22). While this may have enhanced methodological homogeneity, it could also have increased the reproducibility of positive results if selection bias is important.

Third, the treatment effect sizes shown may vary according to different variables. For example, higher proportions of subjects with esophagitis or predominant heartburn confer higher *a priori* chances of a therapeutic response to PPIs. The impact of this variable on estimates of outcome is certainly large, but we decided not to measure it in a formal metaregression that would have been under-powered given the small number of studies included.

## CONCLUSIONS

Although more data are needed on the efficacy and diagnostic utility of PPI therapy in NCCP, our results suggest that therapy with a PPI can be proposed as a first line approach in patients with chest pain of noncardiac origin. The application of a PPI test to identify chest pain patients with abnormal gastroesophageal acid reflux is moderately supported by the available data.

**Reprint requests and correspondence:** Nicholas J Talley, M.D., Ph.D., Mayo Clinic 200 First Street S.W. Plummer 6-56 Rochester, MN 55905.

*Received October 11, 2004; accepted December 15, 2004.*

**Table 4.** Proton Pump Inhibitor Test Performance Summaries According to Different Definitions of "Positive"

| Criteria for Symptom Response | No. of Studies | Pooled* Sensitivity (%, Range) | Pooled* Specificity (%, Range) | Pooled* Diagnostic OR (95% CI) |
|---|---|---|---|---|
| Investigators chosen | 8 | 84 (0–95) | 76 (60–91) | 13.83 (5.48–34.91) |
| >50% response | 7 | 83 (0–92) | 75 (60–91) | 12.59 (4.54–34.92) |
| Any response | 4 | 46 (10–83) | 72 (31–100) | 3.67 (0.92–14.72) |
| Complete relief | 4 | 81 (0–96) | 65 (31–100) | 7.80 (1.04–58.25) |

OR, Odds ratio.
*Pooled using the DerSimonian–Laird method.

## REFERENCES


1. Eslick GD. Noncardiac chest pain: Epidemiology, natural history, health care seeking, and quality of life. Gastroenterol Clin N Am 2004;33:1–23.
2. Eslick GD, Coulshed DS, Talley NJ. Review article: The burden of illness of non-cardiac chest pain. Aliment Pharmacol Ther 2002;16:1217–23.
3. Fass R, Naliboff B, Higa L, et al. Differential effect of long-term esophageal acid exposure on mechanosensitivity and chemosensitivity in humans. Gastroenterology 1998;115:1363–73.
4. Beedassy A, Katz PO, Gruber A, et al. Prior sensitization of esophageal mucosa by acid reflux predisposes to reflux-induced chest pain. J Clin Gastroenterol 2000;31:121–4.
5. Kahrilas PJ. Esophageal motility disorders: Current concepts of pathogenesis and treatment. Can J Gastroenterol 2000;14:221–31.
6. Richter JE, Barish CF, Castell DO. Abnormal sensory perception in patients with esophageal chest pain. Gastroenterology 1986;91:845–52.
7. Wise CM, Semble EL, Dalton CB. Musculoskeletal chest wall syndromes in patients with noncardiac chest pain: A study of 100 patients. Arch Phys Med Rehabil 1992;73:147–9.
8. Olden K. The psychological aspects of noncardiac chest pain. Gastroenterol Clin N Am 2004;33:61–7.
9. Fass R, Fennerty MB, Ofman JJ, et al. The clinical and economic value of a short course of omeprazole in patients with noncardiac chest pain. Gastroenterology 1998;115:42–9.
10. Ofman JJ, Gralnek IM, Udani J, et al. The cost-effectiveness of the omeprazole test in patients with noncardiac chest pain. Am J Med 1999;107:219–27.
11. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. Lancet 1999;354:1896–900.
12. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.
13. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.
14. Deeks JJ. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ 2001;323:157–62.
15. Achem SR, Kolts BE, MacMath T, et al. Effects of omeprazole versus placebo in treatment of noncardiac chest pain and gastroesophageal reflux. Dig Dis Sci 1997;42:2138–45.
16. Bautista J, Fullerton H, Briseno M, et al. The effect of an empirical trial of high-dose lansoprazole on symptom response of patients with non-cardiac chest pain—A randomized, double-blind, placebo-controlled, crossover trial. Aliment Pharmacol Ther 2004;19:1123–30.
17. Fass R, Fullerton H, Hayden CW, et al. Patients with noncardiac chest pain receiving an empirical trial of high dose rabeprazole demonstrate early symptom response—A double-blind, placebo-controlled trial. Gastroenterology 2002;122:A580–81.
18. Pandak WM, Arezo S, Everett S, et al. Short course of omeprazole: A better first diagnostic approach to non-cardiac chest pain than endoscopy, manometry, or 24-hour esophageal pH monitoring. J Clin Gastroenterol 2002;35:307–14.
19. Squillace SJ, Young MF, Sanowski RA. Single dose omeprazole as a test for noncardiac chest pain. Gastroenterology 1993;104:A197.
20. Xia HH, Lai KC, Lam SK, et al. Symptomatic response to lansoprazole predicts abnormal acid reflux in endoscopy-negative patients with non-cardiac chest pain. Aliment Pharmacol Ther 2003;17:369–77.
21. Chambers J, Cooke R, Anggiansah A, Owen W. Effect of omeprazole in patients with chest pain and normal coronary anatomy: Initial experience. Int J Cardiol 1998;65:51–5.
22. Dekel R, Martinez-Hawthorne SD, Guillen RJ, et al. Evaluation of symptom index in identifying gastroesophageal reflux disease-related noncardiac chest pain. J Clin Gastroenterol 2004;38:24–9.
23. Numans ME, Lau J, de Wit NJ, et al. Short-term treatment with proton-pump inhibitors as a test for gastroesophageal reflux disease: A meta-analysis of diagnostic test characteristics. Ann Intern Med 2004;140:518–27.
24. Veldhuyzen van Zanten SJ, Talley NJ, Bytzer P, et al. Design of treatment trials for functional gastrointestinal disorders. Gut 1999;45 (Suppl 2):II69–77.