Complications of Proton Pump Inhibitor Therapy

Michael F. Vaezi,1 Yu-Xiao Yang,2 and Colin W. Howden3

1Division of Gastroenterology, Hepatology, and Nutrition, Vanderbilt University Medical Center, Nashville, Tennessee; 2Division of Gastroenterology and Center for Clinical Epidemiology and Biostatistics, The Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania; and 3Division of Gastroenterology and Hepatology, University of Tennessee Health Science Center, Memphis, Tennessee

Safety issues associated with proton pump inhibitors (PPIs) have recently attracted widespread media and lay attention. Gastroenterologists are frequently asked about the appropriateness of PPI therapy for specific patients. Furthermore, some patients may have had PPI therapy discontinued abruptly or inappropriately due to safety concerns. Faced with such a wide variety of potentially serious adverse consequences, prescribers need to evaluate the evidence objectively to discern the likelihood that any reported association might actually be causal. Here, we review many of the proposed adverse consequences of PPI therapy and apply established criteria for the determination of causation. We also consider the potential contribution of residual confounding in many of the reported studies. Evidence is inadequate to establish causal relationships between PPI therapy and many of the proposed associations. Residual confounding related to study design and the overextrapolation of quantitatively small estimates of effect size have probably led to much of the current controversy about PPI safety. In turn, this has caused unnecessary concern among patients and prescribers. The benefits of PPI therapy for appropriate indications need to be considered, along with the likelihood of the proposed risks. Patients with a proven indication for a PPI should continue to receive it in the lowest effective dose. PPI dose escalation and continued chronic therapy in those unresponsive to initial empiric therapy is discouraged.

**Keywords:** GERD; PPI; Complications.

Proton pump inhibitors (PPIs) are among the most commonly prescribed medicines for gastroesophageal reflux disease (GERD) and peptic ulcer disease.1 In 2015 in the United States, this class of medication ranked in the top 10 national health-related drug expenditures.2 PPIs block acid production by irreversibly inhibiting H+/K+-adenosine triphosphatase in gastric parietal cells. As such, they are often the treatment of choice for acid-related disorders. Omeprazole, the first drug in this class, was introduced in 1989 and was followed by lansoprazole (1995), rabeprazole (1999), pantoprazole (2000), esomeprazole (2001), and dexlansoprazole (2009). Multiple randomized controlled trials (RCTs) with PPIs have shown efficacy for their Food and Drug Administration–approved indications3 (Table 1). Current guidelines recommend empiric therapy with PPIs for patients suspected of having GERD.4,5 They are generally well tolerated, with rare adverse reactions, including flatulence, headache, diarrhea, abdominal pain, and nausea, which are often self-limiting or can be addressed by switching to a different PPI.

There is growing concern regarding the utilization of PPIs. In the United States, omeprazole, esomeprazole, and lansoprazole are available for over-the-counter purchase, resulting in increased public access. Although over-the-counter PPIs are approved only for the short-term management of frequent heartburn, they are also often used for other upper gastrointestinal symptoms, including abdominal pain, bloating, and belching. Furthermore, PPIs are used off-label for functional dyspepsia and for the long-term management of Barrett’s esophagus. They are commonly prescribed at unapproved twice-daily dosing in patients with extraesophageal manifestations of GERD, or GERD symptoms that have not been adequately controlled with once-daily use. In addition to their increased and sometimes inappropriate use, there are now questions about the potential long-term adverse outcomes associated with PPIs. In response to this growing concern, we review the current evidence on many of these reported associations (Figure 1). Rather than providing an exhaustive overview of all reported adverse consequences of PPI therapy, our aim here is to provide perspective on the likelihood of causality versus association based on available observational studies. We evaluate the current evidence based on the established Hill criteria6 (Table 2) to help health care providers, as well as their patients, better interpret the current publications. It should, however, be noted that there are no hard and fast rules by which to judge causation; we have applied the Hill criteria to help formulate an organized approach to making reasonable judgments about the evidence. In 1965, Sir Austin Bradford Hill proposed a list of 9 considerations (Table 2) to strengthen the notion of causality.

Abbreviations used in this paper: AIN, acute interstitial nephritis; BMD, bone mineral density; CAP, community-acquired pneumonia; CDI, Clostridium difficile infection; CI, confidence interval; GERD, gastroesophageal reflux disease; HE, hepatic encephalopathy; H2RA, histamine H2-receptor antagonists; NSAID, nonsteroidal anti-inflammatory drug; OR, odds ratio; PA, pernicious anemia; PPI, proton pump inhibitor; RCT, randomized controlled trial; RR, relative risk.

Most current article

© 2017 by the AGA Institute
0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2017.04.047
versus association. This is important because most epidemiologic studies suggest only associations and as such are prone to various biases leading to erroneous extrapolation to causality. For example, elderly patients may be at increased risk of adverse effects from PPIs, just as they are from some other medicines. The elderly are more likely to have comorbid conditions and to be prescribed aspirin or nonsteroidal anti-inflammatory drugs (NSAIDs). In such a patient, PPI therapy might be entirely appropriate but yet may be blamed for any subsequent adverse event to which the patient had already been predisposed—an example of channeling bias.7

### False Alarms and the Hill Criteria

The current evidence regarding associations of PPI use with adverse long-term outcomes is predominantly based on observational studies. Such epidemiologic studies often trigger “false alarms.”8 Reported associations may be false due to inappropriate design or confounding due to poorly adjusted study parameters applied to retrospective analyses. Overzealous conclusions based on weak associations may result in “epidemics” of sensationalized news coverage followed by widespread patient alarm.8 Potentially even more important, however, is that this may lead to inappropriate discontinuation of a medicine that is needed for an established disease process. Alternatively, the finding of an association that is based on well-established criteria can warn about an undesirable and preventable outcome. However, the studies of association between chronic PPI use and various outcomes (Figure 1) have indeed resulted in an awareness not previously appreciated and, as such, have appropriately questioned the overutilization of PPIs for nonapproved indications. Furthermore, these concerns have forced providers to question the long-term use of these agents in patients who may not need chronic therapy. Of importance, the question “What to do for a patient who truly needs chronic PPI therapy?” is one that can be addressed only by thoughtful review of the data that is based on more than associations but that might help to establish true causality.

### Strength of Association

Assessing the strength of association is critical in causality evaluation. It is arguably the most important criterion in evaluating data from observational studies. Because most adverse outcomes are multifactorial conditions, it is not surprising that the reported relationships between PPI therapy and adverse outcomes are quite modest, particularly on the absolute scale9 (Table 3). Weaker associations are less likely to be causal because, for associations that could be explained by some other factor (eg, residual confounding), the effect of the extraneous factor would have to be larger for strong than for weak associations. In general, extraneous factors with a stronger effect on the outcome should be more evident than weaker ones. This is a particularly important consideration when evaluating the evidence base regarding PPI-related adverse effects, which consists mainly of observational studies and is thus susceptible to undetected biases, particularly residual confounding. As such, the causal inference from observational studies must be guarded if the reported effect estimates range from an odds ratio (OR) of 0.33 to 3, also referred to as the “zone of potential bias”10 (Figure 2).

However, the likelihood of a causal association does not necessarily increase with the strength of association; a strong association also could be due to strong confounding. In addition, in some instances, if the magnitude of the association exceeds what one would expect, it may indicate unmeasured confounding or another source of bias. For example, the concern related to PPI therapy and clopidogrel interaction is based on the notion that a PPI would reduce the antiplatelet effect of clopidogrel through competition for binding sites at CYP2C19.10 As such, the magnitude of potential harm conferred by the interaction of clopidogrel with PPIs is inherently limited by the magnitude of cardiovascular benefit conferred by clopidogrel. Depending on the indication for clopidogrel, this benefit is generally modest with relative risk (RR) reductions of 10% to 30%.11–13 Therefore, studies that reported larger effect estimates13 for concomitant use of PPIs should actually raise suspicion about the validity of the findings.

Furthermore, an association that is weak does not preclude a causal relationship and does not necessarily imply a lack of clinical importance. For individual PPI users, the portion of their risk of adverse effects that is attributable to PPI use is quite marginal, even if the associations are proven to be causal. However, it is important to consider the impact of an exposure at the population level. A commonly used disease measure to quantify this impact at the population level is the population attributable risk, which would be calculated by multiplying the attributable risk associated with PPI use by the prevalence of PPI exposure in the population. Because of the high prevalence of PPI therapy in the population, even a relatively modest association with a clinically important outcome (eg, myocardial infarction, hip fracture) would have important public health implications.

### Consistency

Some of the proposed associations with PPI use have not been consistently demonstrated. Among the reported adverse events associated with long-term PPI use, the possible increased risk of fracture has attracted widespread attention. A meta-analysis of 10 studies13 reported a pooled OR for hip fracture associated with PPI use of 1.25 (95% confidence interval [CI] 1.14–1.37). Six studies had demonstrated a positive association with hip fracture (all

---

**Table 1. Food and Drug Administration Indications for PPI Use**

<table>
<thead>
<tr>
<th>Indication</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Healing of erosive esophagitis</td>
</tr>
<tr>
<td>- Maintenance of healed erosive esophagitis</td>
</tr>
<tr>
<td>- Treatment of GERD</td>
</tr>
<tr>
<td>- Risk reduction for gastric ulcer associated with NSAIDs</td>
</tr>
<tr>
<td>- Helicobacter pylori eradication to reduce the risk of duodenal ulcer recurrence in combination with antibiotics</td>
</tr>
<tr>
<td>- Hypersecretory conditions including Zollinger-Ellison syndrome</td>
</tr>
<tr>
<td>- Short-term and maintenance treatment of duodenal ulcer</td>
</tr>
</tbody>
</table>

---
Figure 1. Schematic diagram of some of the reported adverse consequences of long-term PPI therapy.
with ORs <2), and the remaining 4 had shown no significant association, 2 of which actually demonstrated lower fracture incidence among PPI users than controls. Three of the 4 cohort studies had not shown a significant association, whereas 5 of 6 case-control studies had; all of the case-control studies had quantitatively small ORs between 1.20 and 1.62. In general, higher-quality studies have produced lower estimates of risk than lower-quality studies.

Similarly, for community-acquired pneumonia (CAP), the association was initially suggested by a retrospective study conducted among patients with GERD. However, this was not subsequently confirmed in a study examining PPI use and CAP among patients using NSAIDs. Subsequent studies have shown no association between PPI use and CAP.

There has been similar inconsistency among reports linking PPI use to increased risk of spontaneous bacterial peritonitis in patients with cirrhosis. Although 2 meta-analyses had reported significant associations, these were largely based on reports that had not been fully peer-reviewed. Subsequent large, prospective studies have failed to show a significant association.

Regarding the possible association between PPI use and bacterial enteric infections, a meta-analysis of more than 10,000 patients had reported a pooled OR of 3.33. In this instance, consistency has been shown among various studies examining this association. Four studies had reported ORs for the association between PPI use and Salmonella infection ranging from 1.6 to 8.3. For Campylobacter, 6 studies had reported significant ORs ranging from 1.6 to 11.7. This makes biological sense because reduction in gastric acidity impairs one of the body’s natural defense mechanisms against ingested microorganisms. Aside from bacterial enteric infections, PPI use may be associated with diarrhea from other mechanisms, including microscopic colitis, as has been consistently demonstrated as a class effect.

### Table 2. Hill Criteria

<table>
<thead>
<tr>
<th>Strength of association</th>
<th>Consistency</th>
<th>Temporality</th>
<th>Biological gradient</th>
<th>Biological plausibility</th>
<th>Coherence</th>
<th>Experiment</th>
<th>Analogy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the association of high magnitude?</td>
<td>Are the findings reproducible?</td>
<td>Does the use of PPIs precede the observed outcome?</td>
<td>Is there a direct relationship between dose or duration of PPI use and the outcome?</td>
<td>Are there a rational and theoretical basis for the proposed association?</td>
<td>Any conflicts with what is known about the natural history and biology of the disease?</td>
<td>Are the data based on experiments?</td>
<td>Are there features of association similar to other associations judged to be causal?</td>
</tr>
</tbody>
</table>

### Table 3. Absolute and RRs for Adverse Effects Associated With Long-Term PPIs

<table>
<thead>
<tr>
<th>Potential Adverse Effect</th>
<th>Relative Risk</th>
<th>Reference for Risk Estimate</th>
<th>Reference for Incidence Estimate</th>
<th>Absolute Excess Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic kidney disease</td>
<td>10% to 20% increase</td>
<td>Lazarus et al (18)</td>
<td>Lazarus et al (18)</td>
<td>0.1% to 0.3% per patient/year</td>
</tr>
<tr>
<td>Dementia</td>
<td>4% to 80% increase</td>
<td>Haenisch et al (90)</td>
<td>Haenisch et al (90)</td>
<td>0.07% to 1.5% per patient/year</td>
</tr>
<tr>
<td>Bone fracture</td>
<td>30% to 4-fold increase</td>
<td>Yang et al (27)</td>
<td>Yang et al (27)</td>
<td>0.1% to 0.5% per patient/year</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>No association in RCTs</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Small intestinal bacterial overgrowth</td>
<td>2-fold to 8-fold increase</td>
<td>Lo et al (31)</td>
<td>None available</td>
<td>Unable to calculate</td>
</tr>
<tr>
<td>Campylobacter or Salmonella infection</td>
<td>2-fold to 6-fold increase</td>
<td>Bavishi et al (26)</td>
<td>Crim et al (92)</td>
<td>0.03% to 0.2% per patient/year</td>
</tr>
<tr>
<td>Spontaneous bacterial peritonitis</td>
<td>50% to 3-fold increase</td>
<td>Xu et al (33)</td>
<td>Fernandez et al (94)</td>
<td>3% to 16% per patient/year</td>
</tr>
<tr>
<td>Clostridium difficile infection</td>
<td>No risk to 3-fold increase</td>
<td>Furuya et al (95)</td>
<td>Lessa et al (96)</td>
<td>0% to .09% per patient/year</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>No association in RCTs</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Micronutrient deficiencies</td>
<td>60% to 70% increase</td>
<td>Lam et al (97)</td>
<td>Bailey et al (98)</td>
<td>0.3% to 0.4% per patient/year</td>
</tr>
<tr>
<td>Gastrointestinal malignancies</td>
<td>No association in RCTs</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

**NOTE.** This table provides absolute and RR estimates based on RCTs, meta-analyses, or large observational studies. The purpose of this table is to enable easy comparison of absolute and RRs. Readers should not assume that we believe there is causal relationship when risk estimates are given; Table 3 provides our best summary of the evidence for potential PPI-associated adverse effects.

*Estimates are for adults (mean age 50 years) with a baseline estimated glomerular filtration rate >60 mL/min/1.73m².

*Estimates are for noninstitutionalized adults age 75 years or older.

*Estimates are for adults with a mean age of 77 years.

*Estimates are for patients with cirrhosis with ascites and assume use of spontaneous bacterial peritonitis prophylaxis with antibiotics.

*Estimates are for community-acquired CDI.

*Estimates are for noninstitutionalized adults and are based on vitamin B₁₂ deficiency, defined by both a low vitamin B12 level and an elevated methylmalonic acid level.
However, there are some important caveats about consistency. First, consistency among studies carries more weight if the studies used different designs and patient groups, and still arrived at the same conclusion. Conversely, if all the studies had used the same methodology, they could just have consistently replicated the same inherent bias. Another issue is that apparent consistency in a body of evidence could reflect publication bias.

**Specificity**

The specificity criterion has limited utility because so many conditions are of multifactorial etiology. Furthermore, it has generally not been fulfilled for the suggested associations with PPI use. For example, there are many possible etiological or predisposing factors for both hip fracture and CAP; it is likely that the association between PPI use and hip fracture risk may simply have been a result of confounding, and its significance has been overinterpreted. In the initial study examining PPI use and hip fracture, associations were also noted between hip fracture and other drug classes, including anxiolytics and anticonvulsants. Intuitively, these make biological sense because patients who are sedated or who have seizure disorders are clearly at risk of falls and subsequent trauma, including hip fractures. A study from the Kaiser Permanente group had shown that the increased incidence of fracture among PPI users was evident only in those with other risk factors for fracture.

For CAP, PPI use may simply have been a surrogate for underlying GERD, which is a known predisposing factor. When the risk factor of GERD was circumvented by studying the rate of CAP among people taking PPIs because of NSAID use, the effect was not seen.

Hypomagnesemia and rhabdomyolysis have both been associated with PPI use. However, no specificity has been demonstrated in various sporadic case reports and observational studies. In other more credible predisposing factors have been evident, such as diuretic use and hereditary predisposition in some patients with hypomagnesemia and statin use in patients with rhabdomyolysis. Individuals with congenital mutations in transmembrane receptor potential melastatin protein channels may be predisposed to hypomagnesemia that becomes apparent only with PPI use.

Some of the adverse events attributed to PPIs are idiosyncratic in nature and, therefore, nonspecific. Examples would include acute interstitial nephritis (AIN) and subacute cutaneous lupus erythematosus.

**Temporality**

Most published reports on PPI safety issues are case-control studies; cohort studies and RCTs have been reported much less often. Only incident outcome events are considered in these study designs. In addition, for most potential adverse effects, there is a long induction period between the PPI exposure and outcome. Therefore, the mechanistically relevant exposure has generally been intermediate- to long-term PPI therapy. For these reasons, temporality (ie, PPI exposure preceding the adverse effect) has generally been satisfied. However, there are 2 important issues that warrant consideration.

The first is protopathic bias (Table 4), which occurs when a drug is used to treat early signs of the outcome, giving the appearance that the drug is causally associated with the outcome; conceptually it is analogous to reverse causality. This bias was probably present in studies regarding the association between PPI therapy and CAP. In initial studies reporting this association, there was an unusual yet overlooked trend. Specifically, the increased CAP risk was most pronounced among current PPI users who started PPI therapy within the preceding 14 or 30 days, and the risk increase was attenuated or nonexistent among those current users who had started a PPI in the more remote past. This is contrary to what one would expect if PPIs increased the risk for CAP by induction of hypochlorhydria or through immunosuppression. Sarkar et al subsequently elucidated this issue in a study conducted within a UK general practice database. They first demonstrated that any increase in CAP risk was restricted to patients who started PPI therapy within 30 days. In addition, they showed that the risk of CAP was progressively larger among recipients who started PPI therapy within 14 (adjusted OR, 3.16; 95% CI, 2.45–4.08), 7 (adjusted OR, 3.80; 95% CI, 2.70–5.41), or even 2 (adjusted OR, 6.53; 95% CI, 3.95–10.80) days before the index date. If such a dramatic risk increase among recipients of PPI therapy for 2 days or less had a biological basis, it should have been more pronounced, or at least persistent, among those who had been receiving the medication for longer periods. Furthermore, there was a similar inverse temporal trend of increase in the risk of CAP among new recipients of histamine H2-receptor antagonists (H2RAs) at the same or even slightly larger magnitude. Because H2RAs are weaker acid suppressants than PPIs, these results provided further evidence that the increased risk for CAP seen among new PPI recipients was unlikely to have resulted from the acid-suppressive effect of the PPIs. Based on these considerations, Sarkar et al proposed that these findings may be more consistent with protopathic bias. Specifically, early symptoms related to ensuing CAP, such as cough or chest discomfort, might have been mistaken for acid-related symptoms and treated empirically with PPIs shortly before the eventual clinical diagnosis of CAP. Another possible scenario is that PPIs could have been initiated for symptoms caused by NSAID therapy started for early symptoms of CAP.

Another example of protopathic bias in the PPI safety literature is the positive association between recent (<1 year) initiation of PPI therapy and an increased risk of
Residual confounding This bias occurs when there is persistence of a portion of the confounding effect of a measured confounder.

Biological Gradient

A gradient effect refers to the presence of a monotonic dose- or duration-response relationship between the exposure and outcome. The presence of such an effect has not been consistently demonstrated for many of the PPI safety issues. For example, regarding the association between PPI therapy and fracture risk, among 5 studies that reported effect estimates by dose,27,28,40–42 4 reported a stronger effect with the higher-dose PPI therapy. However, there was significant heterogeneity with respect to the magnitude of the dose-related differences and how low-versus high-dose was defined. Among 9 studies that reported both short- and long-duration effects,27,28,40–45 only 427,28,43,44 observed a duration-response effect. Again, there was significant heterogeneity with respect to the magnitude of the duration-related differences and how short- versus long-duration therapy was defined. A meta-analysis on PPI use and bone fracture found neither a dose-response nor a duration-response effect for PPI use and fracture risk, although substantial heterogeneity among studies made it difficult to interpret the pooled effect estimates.15 Similar inconsistencies have been observed among studies for other PPI–adverse effect associations.

There are several important caveats regarding the gradient effect in this context. First, most of the data sources are not life-time databases and cannot capture PPI use before a patient was included in the database. Therefore, there could be a variable amount of misclassification (ie, underascertainment in the current context) in the duration of PPI therapy, complicating the interpretation of duration-response analysis.27 Second, some causal associations may be characterized by a threshold effect rather than a monotonic trend. For example, osteoporotic fractures generally do not occur until bone mineral density (BMD) drops below the fracture threshold. Therefore, a potential detrimental effect of PPI on fracture risk might not follow a simple monotonic dose-response effect. Third, a monotonic trend with increasing levels of exposure is not necessarily causal. In fact, it could result from a confounding factor that demonstrates a biological gradient in its relation to the adverse outcome. For example, a dose-response effect between PPI therapy and hip fracture risk could reflect the progressive relationship between the gradation of poor general health status and the hip fracture risk.

Plausibility and Experiment

A biologically plausible explanation for any proposed association provides a rational and theoretical basis for the link between the proposed exposure and the observed outcome. The overarching biological explanations proposed for the adverse outcomes linked to chronic PPI therapy have generally been based on gastric acid suppression or idiosyncratic effects of these agents (Table 5). Many are based on data acquired only through animal models or tissue culture with questionable clinical relevance. Below, we examine the proposed biological plausibility for some of the most prominent associated outcomes.

Kidneys: Acute and chronic kidney disease. One proposed mechanism for acute and chronic kidney disease involves a rare but probably idiosyncratic effect of PPIs on the
Table 5. Proposed Mechanisms of Chronic Complications of PPI Therapy

<table>
<thead>
<tr>
<th>Organ</th>
<th>Mechanism</th>
<th>Proposed Effect</th>
</tr>
</thead>
</table>
| Kidney | Recurrent AIN | a) Decreased gastric acidity leading to vitamin B12 deficiency  
| | | b) Beta-amyloid deposition |
| Brain | | a) Decreased gastric acidity leading to reduced calcium and vitamin B12 absorption |
| | | b) Hypergastrinemia leading to hyperparathyroidism |
| Bone | | a) Decreased gastric acidity leading to reduced calcium and vitamin B12 absorption |
| | | b) Hypergastrinemia leading to hyperparathyroidism |
| Heart | | a) Inhibiting clopidogrel activation (Cytochrome P2C19)  
| | | b) Increased asymmetric dimethylarginine leading to reduced endothelial nitrous oxide resulting in thrombosis |
| Colon | | a) Decreased gastric acidity altering intestinal normal flora |
| | | b) Trophic effect of hypergastrinemia on colonocytes |
| Lungs | | a) Decreased gastric acidity and overgrowth of gastric bacteria |
| | | b) Antineutrophilic effect of PPIs |
| Muscle | CYP3A4 enzyme inhibition |
| Blood | Decreased gastric acidity leading to iron and vitamin B12 deficiencies |
| Liver | a) Altered gut microbiota due to gastric acid suppression  
| | | b) Vitamin B12 deficiency due to reduced gastric acid |
| Stomach | Acid suppression induced parietal cell hyperplasia |

Kidneys leading to recurrent AIN. AIN is a humoral- and cell-mediated hypersensitivity reaction resulting in inflammation of the renal interstitium and tubules. PPI-induced AIN may be subtle and without systemic allergic manifestations. The elderly may be at particular risk. The first report of omeprazole-induced AIN was in 1992; more cases of AIN have been reported subsequently. Recent epidemiologic studies have suggested an association between PPI use and chronic kidney disease due to recurrent episodes of AIN.

**Brain: Dementia.** The biological rationale proposed for PPI use and dementia depends on 2 proposed mechanisms, namely vitamin B12 deficiency and enhanced brain beta-amyloid levels. Vitamin B12 deficiency has been associated with cognitive decline; previous epidemiologic studies had suggested a link between gastric acid suppression and reduced vitamin B12 levels. PPI-induced hypochlorhydria may lead to impaired release of dietary protein-bound vitamin B12 that can be absorbed from the terminal ileum. Recent animal data have suggested that some PPIs may interact with brain enzymes. In studies in a mouse amyloid model, Badiola et al observed increased beta-amyloid levels possibly due to inverse gamma secretase BACE1 activity. An alternative explanation that has been proposed for increased beta-amyloid levels is through decreased degradation by lysosomes in microglia. This process may be due to inhibition of vacuolar type H+-ATPase leading to increased pH and reduced clearance of beta-amyloid peptides, which are a major pathologic feature of dementia in Alzheimer disease.

**Bone: Fracture and osteoporosis.** One proposed mechanism for bone loss with subsequent fracture risk and PPI use involves reduction of gastric acidity with subsequent hypergastrinemia. The former may lead to malabsorption of calcium and vitamin B12 and the latter to secondary hyperparathyroidism. Vitamin B12 deficiency may lead to homocysteinemia linked to reduced bone strength. In animal models, hypergastrinemia due to PPI therapy resulted in parathyroid hyperplasia leading to reduced femur bone density. In addition, limited evidence suggests that PPIs might theoretically inhibit osteoclastic vacuolar proton pumps, leading to reduced bone resorption. However, the concentration of omeprazole required to act on the vacuolar H+-ATPase in osteoclasts is approximately 800-times higher than that required to inhibit parietal cell H+/K+-ATPase. PPIs administered orally are unlikely to reach such high concentrations in the subcellular sealing zone. Therefore, an effect on osteoclastic proton pumps is unlikely to be physiologically relevant.

**Heart: Myocardial infarction.** PPIs may compete with the hepatic cytochrome P450 2C19 isoenzyme, thereby inhibiting clopidogrel activation in those with acute coronary syndrome. This would increase the likelihood of clot formation in patients at risk for coronary thrombosis and myocardial infarction. An alternative proposed mechanism invokes reduction of endothelial nitrous oxide through PPI inhibition of dimethylarginine dimethylaminohydrolase enzymatic activity responsible for clearance of asymmetric dimethylarginine, thereby reducing nitrous oxide synthase activity. Similar to many of the associations and biological plausibility in this area, these findings have largely been limited to ex vivo studies.

**Colon: Clostridium difficile infection and microscopic colitis.** Reduction of gastric acidity by PPIs could influence the composition of the normal intestinal flora. Although C difficile spores are not affected by gastric acidity, it is proposed that vegetative forms that would normally be destroyed by gastric acid might survive in the less-acidic environment associated with PPI use and hence predispose to infection. However, a more recent in vitro experiment has challenged these postulated biological mechanisms. PPIs are among the classes of drugs to have been causally linked to microscopic colitis. In a population-based case-control study in the Netherlands, the OR for microscopic colitis associated with PPI use was 7.3 (95% CI, 4.5–12.1). The underlying pathophysiology to explain this association is unclear but may be related to PPI-induced changes in the functioning of intercellular tight junctions or alterations in the colonic microbiome.
Lungs: Pneumonia. PPI-induced gastric acid reduction may lead to growth of aerobic bacteria in the stomach, which may subsequently lead to micro-aspiration and lung colonization with the potential of causing pneumonia. However, this finding is based solely on in vitro studies and its relevance is uncertain.

Muscle: Myopathy. Co-administration of a PPI with an NSAID or statin has been reported to result in rhabdomyolysis. Statins are usually metabolized by the CYP3A4 and PPIs can inhibit this enzyme to a variable extent. Thus, the proposed mechanism involves inhibition of statin metabolism by PPI therapy, leading to dose-related adverse effects, including myopathy.

Blood: Anemia. The predominant mechanism for anemia in patients on PPI therapy is through reduction in gastric acidity that is proposed to lead to reduced absorption of iron and vitamin B12. When gastric pH is greater than 3, ferric ions are poorly reduced to ferrous ions, leading to decreased absorption. Cobalamin is a water-soluble vitamin highly bound to dietary protein. In normal gastric acidity states, hydrochloric acid and pepsin act to release cobalamin, allowing it to bind to salivary R proteins transferring cobalamin to intrinsic factor. The cobalamin–intrinsic factor complex then allows absorption of cobalamin in the terminal ileum. Hypochlorhydria induced by PPIs is purported to interfere with this absorption process, leading to anemia.

Liver: Hepatic encephalopathy. Two mechanisms have been proposed for increased risk of hepatic encephalopathy (HE) and PPI use. The first requires alternation of gut microbiota; hypochlorhydria induced by PPI use may lead to small bowel bacterial overgrowth that might subsequently be important in the development of minimal and overt HE. The second mechanism involves reduction of vitamin B12 due to reduced gastric acidity. Low levels of vitamin B12 may be predictive of the occurrence of HE after liver transplantation.

Stomach: Fundic gland polyps. Acid suppression is theorized to produce parietal cell hyperplasia, leading to histologic changes and polyposis. PPI use also may predispose to the development of fundic gland cysts and mucous blocking of the fundic pits as a result of reduced flow of glandular secretions.

Residual Confounding

Although not one of the Hill criteria, confounding is arguably the most important extraneous factor that could best explain many of the putative associations between PPI therapy and adverse outcomes. Specifically, the central question is whether the observed positive associations are due to the effects of a PPI or the reasons why it was prescribed (i.e., confounding by indication). Here the main concern is not necessarily the conditions for which the PPI was indicated because they are generally not significantly associated with the outcomes. Rather, the real concern is that PPI users generally have worse overall health status than nonusers. This imbalance has been demonstrated in the study population of virtually all published studies addressing PPI safety concerns. Furthermore, because patients with worse health status are also more likely to develop adverse clinical outcomes, health status could confound the association between PPI therapy and adverse outcomes.

Exception for the descriptive case reports and case series of rare, idiosyncratic reactions (e.g., hypomagnesemia, AIN), virtually all published studies on PPI safety issues used some measures to account for this confounding effect.
Although randomization is the most effective way to address this issue, postmarketing RCTs are rarely feasible due to cost and ethical reasons. One of the very few RCTs (Clopidogrel and the Optimization of Gastrointestinal Events Trial)\(^{81}\) was terminated prematurely due to financial problems of the trial sponsor.

Matching and statistical adjustment were strategies used to control for this effect in most nonrandomized studies. These studies almost invariably relied on diagnostic codes in medical records or claims data to measure comorbidities. This approach often fails to take the severity of the comorbidities into consideration, which may lead to residual confounding. For example, dementia is an important determinant of overall health status and is also a risk factor for falls and fractures; it is thus a potential confounder for the association between PPI therapy and hip fracture. Virtually all published studies on this association adjusted for dementia as a dichotomous variable based on the presence of a diagnostic code for dementia. However, patients with a diagnostic code for dementia could range from those with only mild cognitive impairment to those with complete loss of intellectual and physical capabilities. Existing studies have not captured such gradation of dementia or other similar confounding variables, leaving them susceptible to residual confounding.

**Methodological Guidance for Future Studies**

There has clearly been a substantial rise in the number of published studies on PPI-related adverse effects over the past few years. Despite the large volume, there has been little advance in our understanding of the probable biological rationale for these proposed epidemiological associations, and it has been difficult to develop PPI-prescribing practice guidelines based on the existing data. To address these issues, 2 changes are necessary in this line of research.

First, we need to shift our research effort from simply replicating the gross epidemiological associations and pooling these in meta-analyses to understanding the fundamental biological mechanisms relevant to the associations between PPI therapy and the postulated adverse effects. The latter objective would typically involve prospective studies (ie, RCTs or cohort studies) with the outcome being a mechanistically relevant biomarker or surrogate. For example, more than 30 observational studies have examined the association between PPI therapy and the risk of *Clostridium difficile* infection (CDI) over the past 10 years, with most reporting a positive association. Multiple meta-analyses pooling these data also reached similar conclusions. Despite the consistency in the epidemiological finding, there is hardly any progress in our understanding of the nature of the association. A recent study by Freedberg et al\(^{82}\) examined the effect of omeprazole 40 mg twice daily for 4 weeks on the composition and diversity of human fecal microbiota and relevant metabolomes.\(^{82}\) They observed no change in microbiota diversity or fecal bile acids, but found alterations in taxa associated with CDI, which provide the initial mechanistic basis for the epidemiological observation.

Another recent study by Jackson et al\(^{83}\) also showed that PPIs altered the composition of the human gut microbiota. Future studies elucidating the effect of PPI therapy on the colonization resistance of the colon and small bowel microbiome could further clarify the nature of this association.

Similar to the PPI-CDI association, there also has been an abundance of gross epidemiological data but a dearth of information on relevant mechanisms for the PPI-fracture association. As noted previously, there have been some studies of the effect of PPI therapy on BMD.\(^{74,84}\) Although limited by the use of a convenience sample and crude BMD assessment, these studies represent an important step in the right direction and have provided important insight into this association. Future studies need to focus on the effect of PPI on volumetric BMD and non-BMD–related mechanistic pathways.

Besides a shift in research focus, the design of future studies also should be optimized to improve their usefulness in determining causal inference. Although challenging to perform in the postmarketing setting, whenever possible, the RCT is the design of choice. Ideally, these trials should evaluate clinical outcomes (eg, Clopidogrel and the Optimization of Gastrointestinal Events Trial).\(^{81}\) However, clinical trials evaluating the effect of PPI on intermediate biomarkers (eg, NCT01216293 examining the effect of dexlansoprazole on bone homeostasis) also can be valuable and would probably be more feasible.

The design of observational studies also can be improved in several ways to minimize the effect of confounding and bias. First, if sample size allows, an incident user design is preferred over including both new and prevalent PPI users. The incident user design reduces selection bias, confounding, and adjustment for intermediate characteristics in the causal pathway. It also avoids missing the early effect of PPI therapy on potential adverse effects. Second, to minimize residual confounding, it would be ideal to not only measure the presence or absence of a potential confounder but also to capture the levels of exposure to potential confounders. This generally requires analysis of additional diagnostic and therapeutic data available in the data source. For example, for medication exposure, information on dose, duration, and date of initiating use might be relevant beyond just a never/ever use categorization. For medical illnesses, information on chronicity, severity, and treatment requirement might be pertinent. Third, one might consider the use of novel design or analytical approaches to correct for confounding. For example, using a self-control case series design\(^{85}\) and a prior event rate ratio adjustment method,\(^{86}\) a recent study provided strong evidence that residual confounding probably contributed to the apparent association between PPI therapy and the risk of CAP.\(^{19}\) However, data that would allow us to capture the full effect of the confounder are often absent or very limited. In addition, conditions or assumptions necessary for using novel design and analytical methods are often not met. Therefore, it would be important to perform quantitative assessments of confounding routinely to determine the uncertainty of study results. Such assessments, also known as sensitivity analyses or external...
Table 6. Application of the Hill Criteria to Some of the Proposed Associations With Long-Term PPI Therapy

<table>
<thead>
<tr>
<th>Hill Criteria</th>
<th>Clopidogrel Interaction</th>
<th>Fracture</th>
<th>C. difficile Infection</th>
<th>Bacterial Enteric Infection</th>
<th>Hypomagnesemia (ie, &lt;1.6–1.8 mg/dL)</th>
<th>Severe Hypomagnesemia Syndrome</th>
<th>Rhabdomyolysis</th>
<th>Renal Failure</th>
<th>Dementia</th>
<th>MI</th>
<th>Anemia</th>
<th>HE</th>
<th>FGPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength</td>
<td>Weak</td>
<td>Weak</td>
<td>Weak</td>
<td>Moderate</td>
<td>Weak</td>
<td>N/A</td>
<td>Weak</td>
<td>Weak</td>
<td>Weak</td>
<td>Weak</td>
<td>Weak</td>
<td>Weak</td>
<td>High</td>
</tr>
<tr>
<td>Consistency</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Specificity</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Temporality</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Biological</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Maybe</td>
<td>No</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Plausibility</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Coherence</td>
<td>No(^{d})</td>
<td>No(^{d})</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Experiment</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes(^{e})</td>
<td>No</td>
<td>Yes(^{f})</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Analogy</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes(^{g})</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

FGPs, fundic gland polyps; MI, myocardial infarction; N/A, not assessed; SBP, spontaneous bacterial peritonitis; SCLE, subacute cutaneous lupus erythematosus.

\(^{a}\)Limited to case reports and case series.

\(^{b}\)Lack of temporality probably due to protopathic bias (see text).

\(^{c}\)PPIs carry higher risk than H2RAs; high-dose PPIs carry higher risk than standard-dose PPIs.

\(^{d}\)Conclusions of in vitro and pharmacokinetic studies not replicated in clinical studies.

\(^{e}\)Studies of BMD do not show tendency to osteoporosis during PPI therapy.

\(^{f}\)Reports indicate resolution of hypomagnesemia following PPI withdrawal and positive PPI rechallenge.

\(^{g}\)Reports indicate resolution of AIN following PPI withdrawal.

\(^{h}\)Strong and consistent association between PPI use and development of FGPs in the absence of other conditions known to be associated with FGPs (eg, familial adenomatous polyposis).

\(^{i}\)PPI use might alter iron or B\(_{12}\) absorption.
adjustment, will demonstrate combinations of 3 measures (ie, how strong a confounder would have to be associated with the outcome and the exposure, and how prevalent it would need to be in the population) for confounding to account fully for the observed association. The plausibility of the combinations would help to evaluate the likelihood that the primary association is due to confounding. Sensitivity analyses can and should be performed in virtually all observational studies.

The application of meta-analysis to observational studies is controversial because the high risk of bias/confounding in the individual studies makes the calculation of a single summary effect estimate potentially misleading. Because the PPI safety literature consists predominantly of observational studies, published meta-analyses have focused on reporting pooled summary estimates and drawing conclusions from these. As a result, besides providing a more precise but potentially biased effect estimate, these meta-analyses have offered little value in addressing key questions regarding the causal nature of the reported associations. Future systematic reviews and meta-analyses in this line of research should evaluate the trustworthiness of the individual studies in terms of the presence of bias/confounding and the search for sources of between-study heterogeneity rather than simply calculating summary estimates.

Conclusions

Despite the recent alarm generated by some high-profile, newsworthy publications, few (if any) of the recent scare stories have attempted to balance the benefits of PPIs with their alleged risks. PPIs have revolutionized the management of GERD and have been of enormous benefit to patients who are at risk of upper gastrointestinal ulceration and bleeding from aspirin or NSAIDs. Many recent publications have, however, pointed out that a great number of patients are receiving PPIs unnecessarily for conditions or symptoms for which they would not have been expected to provide benefit. Furthermore, many patients who are on PPI treatment for appropriate indications are receiving excessively high daily doses. It is a sound adage of medical practice that all drugs should be given in the lowest effective dose and for the shortest possible time. This is as true for PPIs as it is for any other drug class.

Given the widespread use of the PPIs and our enormous, worldwide cumulative experience with them, it is reasonable to take a step back for a moment and consider the current situation. Virtually the entire evidence base regarding PPI-related safety concerns consists of observational studies. We need to have a clear understanding of the meaning of a “statistically significant” but modest association from such studies. Statistical significance only takes random errors related to sample size into consideration; it ignores systematic errors. Observational studies, no matter how well performed, may be inherently incapable of accurately discerning weak associations from null effect due to their susceptibility to systematic errors of bias/confounding and other methodological weaknesses. We would be straying beyond the limits of these studies if we made causal inferences based on these data or let such inferences determine clinical practice. Criteria for causation have generally not been made for most of the proposed associations; our interpretations based on the Hill criteria are summarized in Table 6.

Therefore, we advise a pragmatic, “common-sense” approach to this issue. Patients with a clear indication for PPI treatment should continue to receive it in the lowest effective dose. When considering the lowest effective dose, it might be helpful to bear in mind that approximately 3% of white and 15% to 20% of Asian individuals have reduced or absent CYP2C19 enzyme activity (“poor metabolizers”); poor CYP2C19 metabolism can enhance the therapeutic effect of PPIs but could theoretically also increase the potential risk of PPI-associated adverse effects. Furthermore, we must accept that, for some patients, treatment may need to be lifelong. However, multiple “false alarms” related to the safety of PPIs could ultimately lead to inappropriate discontinuation of treatment with potentially serious consequences for some patients. Investigators, the press, and, perhaps, even editors of medical and scientific journals should take responsibility to avoid subjecting the public to what Lewis Thomas called an “epidemic of anxiety” causing unintended harm. The media should take a more balanced, critical, and responsible approach in their reporting of epidemiological data so that weak and inconclusive results are not overinterpreted and presented to the lay public as facts. Researchers engaged in investigations on PPI safety issues should devote more effort toward RCTs whenever possible as well as studies that will advance our understanding of the physiological effects of PPI therapy on mechanistically relevant biomarkers. Retrospective observational studies should be pursued based on only biologically plausible hypotheses. The investigators also should use appropriate methodological tools to mitigate the effect of confounding and quantify how robust the observed associations are to potential unmeasured or uncontrolled confounding. Most importantly, they need to understand the limitations of the observational studies and be more skeptical about their own findings from such studies. Much of the current evidence linking PPI use to serious long-term adverse consequences is weak and insubstantial. It should not deter prescribers from using appropriate doses of PPIs for appropriate indications.

References

4. Kahrilas PJ, Shaheen NJ, Vaezi MF. American Gastroenterological Association Institute technical review on


73. Cats A, Schenk BE, Bloemena E, et al. Parietal cell pro-


78. Shah P, Rhim AD, Haynes K, et al. Diagnosis of perni-


95. Furuya-Kanamori L, Stone JC, Clark J, et al. Comorbid-

96. Lessa FC, Mu Y, Bamberg WM, et al. Burden of Clo-

97. Lam JR, Schneider JL, Zhao W, et al. Proton pump inhibitor


Received March 17, 2017. Accepted April 24, 2017.

Reprint requests

Address requests for reprints to: Michael F. Vaezi, MD, PhD, MSc (Epi), Division of Gastroenterology and Hepatology, Center for Swallowing and Esophageal Disorders, 1660 TVC Digestive Disease Center, Vanderbilt University Medical Center, 1301 Medical Center Drive, Nashville, Tennessee 37232-5280. e-mail: Michael.vaezi@vanderbilt.edu; fax: (615) 322-8525.

Conflicts of interest

The authors disclose the following: Dr Vaezi receives research support from Sandhill Scientific. Dr Howden has consultancies with Takeda, Arazel, and Pfizer Consumer Health. Dr Yang has no conflicts to disclose.