

Editorial

Canadian Association of Gastroenterology Statement on the Putative Link Between Proton Pump Inhibitor Treatment and Gastric Cancer after *Helicobacter pylori* Eradication

A recent paper by Cheung et al. (1) reported that long-term use of proton pump inhibitors (PPIs) was associated with an increased risk of gastric cancer “even after” *Helicobacter pylori* (*H. pylori*) eradication. This has created significant concerns among physicians and patients (2). The Canadian Association of Gastroenterology (CAG) has serious reservations about the validity of this study and, hence, considers it important to provide guidance to its members and their patients regarding the study’s methodological limitations and inappropriate conclusions.

WHAT DID THE STUDY BY CHEUNG ET AL. SHOW?

This study was a large retrospective cohort study based on a Hong Kong health database. Adults who had received a prescription of clarithromycin-based triple therapy for *H. pylori* infection between 2003 and 2012 were included. Those who received repeat eradication therapy were excluded (13.4% of otherwise eligible patients). A Cox proportional hazards model was used with propensity score adjustment, which aimed to take patient comorbidities into account. Among the 63,397 people in this cohort, 153 (0.24%) developed gastric cancer after a median follow up of 7.6 years. Following eradication treatment, PPI users (n=3,271; 5.2%) were more likely to be diagnosed with gastric cancer (hazard ratio [HR] 2.44, 95% CI 1.42–4.20), compared with those who did not use PPIs. The authors also reported that there was an association between increasing risk and longer duration of PPIs use, suggesting a dose response. For people using histamine-2 receptor antagonists (H₂RAs; n=21,729; 34.3%), there was no association with gastric cancer (HR 0.72, 95% CI 0.48–1.07). The authors concluded that “long-term use of PPIs was still associated with an increased GC risk in subjects even after *H. pylori* eradication therapy”. A maxim of epidemiology is that association is not causation and, while the authors acknowledged this possibility, the implication from the article was that ongoing PPI use had

caused gastric cancer in those receiving *H. pylori* eradication therapy.

In our opinion, the conclusions of Cheung et al. are not robust and we outline our main concerns with this article.

INADEQUATE ADJUSTMENT FOR CONFOUNDERS

Inadequate adjustment for confounders was the most serious limitation of the study. In association studies, the unadjusted results are expected to show a positive association between PPI use and gastric cancer. This is due to confounding, since PPI users have been consistently shown to be older, frailer and more likely to have risk factors for gastric cancer (smoking, excess alcohol, obesity, etc.) compared with nonusers (3). Adjustment for these confounders always attenuates substantially the magnitude of the association (2). However, in this study, the association only changed slightly and remained statistically significant after an apparently thorough adjustment for confounders. We believe that this is due to inability to adjust for important unmeasured and inaccurately measured confounders. Adjustment for confounders is an essential step for large database observational studies. Confounders are factors that (a) are associated with both the exposure of interest (e.g., PPI use) and the outcome (e.g., gastric cancer), (b) are distributed unequally among the groups being compared, and (c) are not an intermediary step in the causal pathway from the exposure of interest to the outcome. For the study by Cheung et al., factors such as smoking, alcohol consumption, obesity, age and various comorbidities were dealt with as potential confounders. At first look, the adjustment that was performed in this study appears to have been state-of-the-art: propensity scores were used as adjustment variables in a Cox proportional hazards model. However, the devil is in the details. Propensity methods can only improve adjustment for *measured* confounders: they cannot adjust for unmeasured confounders and will not perform well if the confounders have been inaccurately measured (4). This is precisely

the problem in this cohort study. Administrative databases are not created to answer a specific research question, and thus, they never have precise information on all known confounders. This is the case in the study by Cheung et al., in which data on the most important confounder—namely the persistence of *H. pylori* infection—could not be captured among the included patients, while other important confounders, such as smoking, alcohol use and obesity, were measured inaccurately.

In regard to the determination of *H. pylori* status in this study, patients who did not receive a second course of eradication therapy were considered to be *H. pylori* negative (i.e., eradicated). This is seriously inaccurate and would strongly bias the results towards an association between PPI use and gastric cancer. It is inevitable that some patients in the study cohort had persistent *H. pylori* infection (e.g., did not take or complete the eradication treatment; not tested for *H. pylori* post-treatment; lost to follow-up; declined repeat treatment after documented failed eradication). Cheung et al. downplayed the importance of this limitation, suggesting that only a small proportion of the patients in the cohort would have had persistent *H. pylori* infection. However, persistence of *H. pylori* infection is by far the strongest confounding factor in this study. It is the strongest known risk factor for gastric cancer and is likely to be more common in the PPI group because (a) people with persistent infection are more likely to remain symptomatic (5) and therefore require PPIs and (b) people who failed to stop PPIs prior to post-treatment testing for *H. pylori* were more likely to have false-negative results. This misclassification cannot be considered nondifferential: it systematically deviates the results towards a positive PPI–gastric cancer association. It is the “elephant in the room”, the critically important confounder that is present but eludes measurement. Even if a small proportion of the included patients had persistent *H. pylori* infection, this would correspond to an absolute number large enough to seriously skew the results and invalidate the conclusions of the study. For example, 2% persistent infection translates to 1,268 patients at high risk for developing gastric cancer; this number is large enough to affect the validity of the 134 versus 19 cases of gastric cancer observed in the study.

Another important unmeasured confounder was baseline gastric histology. Results should have been adjusted for baseline presence of gastric atrophy, intestinal metaplasia and dysplasia.

Regarding measured confounders, some degree of inaccuracy is inevitable in all administrative database studies, but in this cohort study, the inaccuracy is more serious than usual. This is readily evident from the table of patient characteristics which shows that the prevalence of some confounding factors was implausibly low: smoking 2.6%, alcohol consumption 0.9% and obesity 1.0%. Obviously, the true prevalence was substantially higher (for example, in 2012, the prevalence of smoking among adults in Hong Kong was estimated to be 10.7%) (6). Cheung

et al. acknowledged that “the identification of certain parameters (smoking, alcohol use and obesity) via coding may underestimate their true prevalence, as only patients who had heavy consumption of smoking and alcohol or who were morbidly obese would be coded” (1). However, this is not a minor limitation; it is a critically important flaw. Smokers and obese patients are more likely to have reflux symptoms and therefore more likely to be prescribed PPI therapy. Moreover, smoking (7) and obesity (8) are also strong risk factors for gastric cancer. This misclassification bias will predictably skew results towards a positive association between PPI use and gastric cancer because the majority of those who smoked, consumed alcohol or were obese escaped adjustment for those strong confounders.

Obviously, the concerns noted above also apply to the inaccuracies in the measurement of other confounding factors, especially comorbidities and dyspepsia. Comorbidities cannot be captured with adequate granularity/gradation in administrative databases; for adequate adjustment, the severity of comorbidities has to be measured as continuous or at least as an ordinal variable. The dichotomous (present versus absent) approach that was used in this cohort study has inevitably allowed for residual confounding. Similarly, it is very unlikely that dyspepsia was captured accurately. There is no code for dyspepsia among the ICD-9 codes that were used. Furthermore, dyspepsia is notoriously difficult to capture in the Chinese language (9).

The end result of not capturing or inaccurately measuring confounding factors in this cohort study was the surprisingly small change of the HRs after the application of an apparently rigorous adjustment for confounders.

Failure to capture or measure confounding factors accurately in this cohort study has almost certainly led to the surprisingly small change of the HRs that was reported despite the application of an apparently rigorous adjustment for confounders.

NO PROOF OF DOSE-RESPONSE OR DURATION-RESPONSE

Numerical differences in the HR for gastric cancer between different frequencies or durations of PPI use were presented as proof of a dose-response relationship although curiously, no tests for statistical significance were shown. This is misleading since statistical testing, which is easily performed, shows no statistically significant differences or evidence of a dose-response effect; in fact, this could have been predicted in view of the fact that the 95% CIs for the different time intervals are wide and clearly overlapping.

SURVEILLANCE BIAS

Cheung et al. acknowledged the possibility of surveillance bias: “PPIs users may have a higher chance to have endoscopy than non-PPIs users resulting in discovery of more gastric cancers

due to surveillance bias” (1). In other words, early gastric cancers were more likely to be diagnosed among PPI users, while in the non-PPI group, such cancers could be diagnosed after the study window or patients might expire from other causes first. The authors argue, “However, as shown in the table 2 in the online supplementary file 1, the incidence rate of gastric cancer remained relatively stable throughout the follow-up period rather than an early peak in the first few years followed by a rapid drop in the ensuing years” (1). In our view, the table results are not incompatible with surveillance bias. The only approach that would provide insight into the magnitude of the surveillance bias would be reporting gastric cancers by stage, but such data were not available.

H₂RA AS NEGATIVE CONTROL EXPOSURE

Negative control exposures (NCEs) can alert us to residual confounding and bias; however, wrong choice and use of NCE, as in this study, can be misleading. The principles for selecting an NCE are simple (10). First, an NCE cannot involve the hypothesized causal mechanism for the main exposure of the study. H₂RAs are not ideal NCEs; although “PPIs are much more potent than H₂RAs in terms of gastric acid suppression”, both of them lead to acid suppression, which is the hypothesized causal mechanism for the PPI–gastric cancer link. Second, NCE cannot be dependent on the main exposure of the study. In this study, it is plausible that there is an inverse (competing) association among H₂RAs and PPIs. Third, the function of an NCE is to challenge the validity of the main results of a study (by showing an implausible result that can only be explained by residual confounding or bias); an NCE cannot further strengthen the main results of a study, as Cheung et al. have attempted to do. The nonsignificant result for the H₂RA–gastric cancer association does not prove or disprove anything about the PPI–gastric cancer association.

ADDITIONAL COHORT OF PPI USERS “TO FURTHER CONTROL FOR POSSIBLE CONFOUNDING EFFECTS”

This additional analysis is seriously misleading. Cheung et al. comment that “by comparing the incidence rate of gastric cancer of a matched cohort of PPIs users who had not received *H. pylori* eradication therapy, we showed that *H. pylori* infection, even prior infection, was a more important factor than PPIs use in determining gastric cancer risk.” (1) This is not correct; the comparison does not help interpret the data; it neither strengthens nor weakens the main result of the study. It is meaningless to compare gastric cancer incidence rates for PPI users who did or did not receive *H. pylori* eradication therapy; the only conclusion from this analysis would be that among PPI users, *H. pylori* eradication therapy is associated with increased risk of gastric cancer—an obviously erroneous conclusion, caused by

confounding (mainly confounding by indication). A meaningful analysis would have been to compare the incidence rates of gastric cancer among *H. pylori*–negative patients who are or are not using PPIs, if such data had been available.

PRECISION VERSUS VALIDITY

This was a large study, but the size of a study only improves precision (reduces the chance or random error), not validity (does not affect systematic errors). A large study with methodological flaws is precisely wrong. If this study was conducted again with 10 times larger sample size, the results would be equally invalid.

CONCLUSIONS

The conclusions of the cohort study by Cheung et al. are unjustified, not only because, as always, “observational studies can only show association and cannot prove causation” but more importantly because it is at high risk of bias because of multiple, critically important flaws inherent in the design and analysis of this particular database study. In summary, the study does not provide any persuasive evidence for a causal link between PPIs and the development of gastric cancer after *H. pylori* eradication therapy. Therefore, physicians should not change practice in their use of PPIs based on the stated conclusions of this study. PPIs should not be withheld from patients who require them; although like all medications, they should be taken at the lowest effective dose and only for as long as clinically indicated.

Conflicts of interest

GIL, SVVZ, LH and NJ have no conflicts to declare. DA has received personal fees and nonfinancial support from Takeda, personal fees and nonfinancial support from Pfizer, personal fees and nonfinancial support from Shire, grants and personal fees from AbbVie, personal fees from Janssen, personal fees from Pendopharm, grants and personal fees from Mylan, personal fees from Allergan and personal fees from Pentax, outside the current work.

PM has received research funding and been on the advisory board from Takeda, has received research funding and been on the advisory board from Allergan, the advisory board from Shire and the advisory board from Lupin.

Grigorios I. Leontiadis MD, PhD^{1,2}, Sander Veldhuyzen Van Zanten MD, FRCPC, MSc, MPH, PhD³, Lawrence Hookey MD, FRCPC^{4,5}, David Armstrong MA, MB BChir, FRCP(UK), FACG, AGAF, FRCPC^{2,6}, Nicola Jones MD, PhD, FRCP(C)^{7,8}, Paul Moayyedi BSc, MB, ChB, PhD, MPH, FRCP, FRCPC, AGAF, FACG^{2,9}

¹VP Clinical Affairs, Canadian Association of Gastroenterology;

²Division of Gastroenterology and Farncombe Family Digestive Health Research Institute, McMaster University, Hamilton, Ontario, Canada;

³Division of Gastroenterology, University of Alberta, Edmonton, Alberta, Canada;

⁴Past VP Practice Affairs, Canadian Association of Gastroenterology;

⁵Division of

Gastroenterology, Queen's University, Kingston, Ontario, Canada;
⁶Past President, Canadian Association of Gastroenterology;
⁷Division of Gastroenterology, Hepatology, and Nutrition, The Hospital for Sick Children, Departments of Paediatrics and Physiology, University of Toronto, Toronto, Ontario, Canada;
⁸President, Canadian Association of Gastroenterology; ⁹VP Quality Affairs, Canadian Association of Gastroenterology
 Correspondence: Grigorios I. Leontiadis MD PhD, VP Clinical Affairs, Canadian Association of Gastroenterology, Associate Professor, McMaster University, Department of Medicine, Division of Gastroenterology and Farncombe Family Digestive Health Research Institute, 1280 Main St. W. HSC 3V3, Hamilton, ON Canada, L8S 4K1, e-mail leontia@mcmaster.ca

References

- Cheung KS, Chan EW, Wong AYS, et al. Long-term proton pump inhibitors and risk of gastric cancer development after treatment for *Helicobacter pylori*: A population-based study. *Gut* 2018;67(1):28–35.
- Moayyedi P, Van Zanten SV, Hookey L, et al. Letter to the editor: Proton pump inhibitors and gastric cancer: association is not causation. *Gut* 2018; in press.
- Moayyedi P, Leontiadis GI. The risks of PPI therapy. *Nat Rev Gastroenterol Hepatol* 2012;9:132–9.
- Agoritsas T, Merglen A, Shah ND, et al. Adjusted analyses in studies addressing therapy and harm: Users' guides to the medical literature. *JAMA* 2017;317:748–759.
- Moayyedi PM, Lacy BE, Andrews CN, et al. ACG and CAG clinical guideline: Management of Dyspepsia. *Am J Gastroenterol* 2017;112:988–1013.
- Hong Kong Council on Smoking and Health. Government released the 2012 Hong Kong Smoking Prevalence. <<http://www.smokefree.hk/en/content/web.do?page=news20131108>> (Accessed March 1, 2018).
- Tredaniel J, Buffetta P, Buiatti E, et al. Tobacco smoke and gastric cancer: Review and meta-analysis. *Int J Cancer* 1997;72:565–73.
- Yang P, Zhou Y, Chen B, et al. Overweight, obesity and gastric cancer risk: Result from a meta-analysis of cohort studies. *Eur J Cancer* 2009;45:2867–73.
- Wang B. Dyspepsia: What you think it is, is different than what doctors think it is [Chinese]. Medical community Gastroenterology Channel. <https://mp.weixin.qq.com/s/UPZKpMcc0_pVs7Yd-q5t08w> (Accessed May 1, 2018).
- Lipsitch M, Tchetgen Tchetgen E, et al. Negative controls: A tool for detecting confounding and bias in observational studies. *Epidemiology* 2010;21:383–8.