Reviewer's report

Title: Association of presence/absence and on/off patterns of Helicobacter pylori oipA gene with peptic ulcer disease and gastric cancer risks: a meta-analysis

Version: 1 Date: 16 April 2013

Reviewer: Yuhong Yuan

Reviewer's report:

This study aimed to study the association of Hp oipA gene and GI diseases PUD and GC, through the method of meta-analysis. As the first meta-analysis on this topic, the study provides some useful information. However, some clarification is required by the authors and some methodological errors are found. Reporting quality is suboptimal and should be improved substantially.

Major comments (Major compulsory revisions)

Controls and definitions:

1. When reporting an odds ratio (OR, defined as the ratio of the odds of an event occurring in one group to the odds of it occurring in another group), reporting the control group is very important. For example, the odds of having the gene in gastric cancer patients compared with gastritis patients may be different than the odds of having the gene in gastric cancer patients compared with healthy volunteers. This report emphasized the “association” by reporting the ORs without providing information on the controls selected. The controls should be provided in the abstract as well as the methods section.

2. Was “superficial gastritis” defined by the primary studies, and was the definition standardized in eligible studies? “Superficial gastritis” is one type of nonerosive gastritis defined by pathology, which is not a commonly used term compared to “chronic gastritis” or “H.pylori associated gastritis” in clinical studies. In addition, it is important to know whether “deep gastritis” and “atrophic gastritis” were excluded from the controls in these studies, while patients with atrophic gastritis are at increased risk for the development of gastric cancer. “Superficial gastritis or NUD” were mixed as the control group in this meta-analysis, other types of gastritis might be found in the “NUD patients” as well.

3. The authors used “NUD” for “non-ulcer disease”, which is my concern. NUD normally refers to “non-ulcer dyspepsia” because “non-ulcer diseases” include many upper GI disease functional as well as organic diseases, such as erosive esophagitis, gastritis, etc. Please clarify whether the review authors meant to include non-ulcer dyspepsia as the control. If so, then according to the new Rome III definition, “functional dyspepsia (FD)” is a more appropriate term than “non-ulcer dyspepsia”. However, if the authors in fact consider any non-ulcer diseases as the controls, then please list all diseases that included in the control groups. The control groups have an important role in interpreting the results.
4. The authors should also justify the combination of gastritis and “NUD” as the controls.

Eligible studies:

5. In the method section, the authors aimed to include studies reported the “association between gene status and clinical outcomes”. However, the inclusion criteria is not clear. First, the study design for eligible study is unclear. Based on the method section, it seems that the authors only considered case-control studies. However, it is unclear whether the authors in fact considered all observational studies (e.g., case-control, cohort, cross-sectional studies) but only case-controls were eligible. Second, whether all cases and controls must be H. pylori positive patients when the studies were assessed (then the authors compared the positive vs negative opiA gene, and the opiA on status on vs off status in strains with positive opiA gene). Third, whether the authors aimed to study any “gastroduodenal diseases” but then only PUD or GC studies met the inclusion criteria, or, only “PUD or GC” was the interested clinical conditions. These terms were used alternative as the outcome throughout the report. It should be clearly stated in the method section. Fifth, it is not clear whether any non-PUD non-GC controls were eligible or only studies included gastritis or NUD as the controls were eligible. Sixth, “studies on human” is redundant (animal studies were excluded). Seventh, excluded “case-only study” should be “case series”; “data covered by other studies” should be “preliminary data of published studies or duplicate publications”.

6. The authors should add “conference proceedings” under the exclusion criteria, because they excluded “meeting abstracts” in the results section.

7. The review authors only searched two English databases (Pubmed and Web of science) and two Chinese databases (CNKI and Wanfang). One of the two most significant health science database Embase was not searched. Embase indexes many journals not covered by PubMed. It should be discussed under the limitations.

Statistical analyses:

8. The report has a “sensitivity analysis” by excluding studies one by one. First, this kind of analysis is not called “sensitivity analysis” in statistics; it should be “influence analysis”. Performing influence analyses is not necessary for this report due to multiple testing. We should only exclude the outlier study as the sensitivity analysis when significant heterogeneity is indicated. The authors should revise the section and clearly state the sensitivity analysis was performed to explore heterogeneity.

9. For four studies (page 9), data from different countries and different age groups were reported but these studies were considered as different studies by the authors. This data handling is only appropriate in the subgroup analyses, not in the overall analyses. The breaking down of the patient groups lead to smaller sample size for the “studies”, increased heterogeneity and the changed of the 95%CI for the point estimate; while the subgroup data reported in the same study
in fact were not different studies. In meta-analyses, different subgroups can have different effect estimates (e.g., gender, age, subtype of disease, etc); considering all subgroups from the same study as different studies in the overall analysis is not justified and introduces analysis bias. We wouldn’t see a proper meta-analysis considering a multinational study as “different studies” in the overall effect estimate. The report and analyses need to be revised accordingly.

10. The authors made some statements about the differences of subgroups. For example, on page 11, “that effects of the presence of oipA gene on PUD development showed difference in different ethnic subpopulations”; and tried to judge the subgroup differences based on their ORs. Whether there are significant subgroup differences, it should be based on the tests for subgroup differences, but not by OR numerically, nor by the overlapping of the 95% CIs. When two 95%CIs are overlapped, the two ORs can be significantly different or can be non-significantly different. When subgroup differences are assessed, the method should be stated. Seeking statistical help is suggested.

11. The authors concluded that no significant publication bias was detected. However, only 7 studies were included in the analysis for the positive gene analysis, therefore, the power is low: “as a rule of thumb, tests for funnel plot asymmetry should be used only when there are at least 10 studies included in the meta-analysis, because when there are fewer studies the power of the tests is too low to distinguish chance from real asymmetry.” (For details, please see Chapter 10 in the Cochrane handbook and Ioannidis JP et al, CMAJ. 2007 Apr 10;176(8):1091-6.). This should be stated under limitations. As discussed, different subgroups should not be considered as independent studies in the publication analysis. Revisions should be made accordingly.

12. Subgroup analyses included “population region and publication language”. Publication language is not a justified subgroup analysis in this report. There is already a subgroup analysis for geographic regions (Asia vs Africa, Europe, and America). If they elect, they can have a subgroup analysis for ethnic groups (e.g., Chinese, Japanese, etc), although it is not as important as the subgroup of countries or regions. However, having a subgroup only based on the published language for Chinese journals vs English journals only (other languages have been excluded) is questionable because English journals publish studies from different regions and mixed different ethnic groups. Unless the authors want to prove the Chinese studies published in English Journals have higher quality than those published in Chinese journals, or Chinese studies have lower quality than other Asian studies, then they can have three subgroups of “Chinese studies in English journals vs Chinese studies in Chinese journals vs non-Chinese studies in English journals”. Furthermore, the authors emphasized in the results and in the discussion that only studies published in Chinese shown positive results. Can it be interpreted that Chinese studies are more likely prone to publication bias? If so, then it should be further discussed. We should not design the subgroup analyses simply based on the convenience or for the purpose of fishing for significant differences. The authors should justify why they need a subgroup analysis based on two languages, when subgroup analyses for geographic regions and ethnic groups are available (please note “studies published in
13. On page 8, the statement “P value > 0.1 for Q test indicates no heterogeneity among studies” is inappropriate, it should be “P < 0.10 indicates significant heterogeneity between studies”. The authors have misconceptions for statistical testing. Just as emphasized in the “The Little Handbook of Statistical Practice” by Dallal: “Null hypothesis are never accepted. We either reject them or fail to reject them. The distinction between “acceptance” and “failure to reject” is best understood in terms of confidence intervals. Failing to reject a hypothesis means a confidence interval contains a value of “no difference”. However, the data may also be consistent with differences of practical importance. Hence, failing to reject H0 does not mean that we have shown that there is no difference (accept H0)”.

14. Similarly, the statement on page 9 is inappropriate: “P value > 0.1 for Begg’s and Egger’s tests indicates no publication bias”.

15. The authors reported data for the GU and DU subgroup, which was not mentioned under the method section-subgroup analyses.

Results

16. On page 11 and figure 2A, the author considered the study Ben as the “most influencing study”, which is not justified. Obviously, the outlier study in this group is “Zhang 2004” with OR as 90.39, while all other studies with OR range from 0.11 to 5.82. As discussed, the authors should consider how to assess outlier study and how to perform sensitivity analysis to explore heterogeneity.

17. GC as an important outcome, only one sentence reported the results of the association between GC and positive oipA gene (on page 11), started with “in addition” and is hidden at the end of the paragraph of GU, DU subgroup. No subgroup or sensitivity analyses were reported although 5 studies were included in this comparison. On the other hand, the authors spent paragraphs and figures for PUD, such as a PUD subgroup of only 2 studies. This indicates this report has high risk of selective reporting. In addition, “(Figure 3A)” should be added at the end of the sentence.

18. The paragraph “correlation between oipA and other virulent factors” should be either removed or revised under discussion. Any data which were not planned a priori in the method section should not be reported in the results section based on authors’ preference (discussed below).

19. OR was “3.91” in Figure 3A but “3.92” in table 2 and on page 11 for PUD and oipA on status.

20. On page 10, “source of H. pylori strain isolates” is reported but not present in supplement table 1 as stated.

21. Table 3, “slight publication” is a subjective expression; it should either be significant, or not significant. Again, using “p > 0.1” for significance is incorrect.

22. Similarly, “slight heterogeneity” is inappropriate on page 12. “Slight one” is inappropriate on page 13.
23. For the readers better understanding of the forest plots, n/N for cases and n/N for controls should be provided in the forest plots for each study. It is very easy to generate from the statistical software.

Results interpretation and discussion:

24. Based on the results, ORs for PUD was higher than GC in patients with “oipA status on gene” (OR 3.92 vs 2.28). Having an increased risk for PUD has a different clinical implication than having an increased risk for GC, especially when considering the cost-benefit of screening and prevention strategies. Do the authors believe that the results imply that the risk of having an ulcer disease is higher than having a cancer in these patients? What is the clinical implication? The authors should improve the results interpretation.

25. As the authors noted, most studies were from Asia. In addition, sample size is small for most of the studies and total event number is small for most comparisons. Studies published in other languages were excluded. The results generalizability therefore is limited based on current evidence, which should be discussed under the limitations.

26. Significant heterogeneity is suggested for most of the comparisons, which could not be explained by the subgroup analyses or sensitivity analyses. It should be discussed under the limitations.

27. On page 14, the authors listed in detailed the number of studies shown association and the number of studies with negative association. This is not necessary as the meta-analysis is performed, and this is what we want: to provide more accurate effect estimate than individual studies.

28. The authors performed different subgroups but did not discuss the clinical implication while the associations were only found for some subgroups but not the other. For example, non-significant association was seen for the GU subgroup but not for DU then the authors explained it as due to “small sample size (n=33) for GU”. On the other hand, when the other subgroup immunoblot was found to be significant although sample size was also small (n=44), the authors avoid discussing the difference with the non-significant subgroup PCR (most commonly used). Instead, a strong conclusion was drawn after different subgroup estimates as “these results indicated that the presence of…” (on page 15).

29. On page 16, the authors “attempted to evaluate the combined effects of other virulence factors”… then stated “One of the main reasons for not performing a combined analysis in those studies may be a limited sample size”, which is misleading. First, the authors did not limit the sample size for the meta-analysis in the method section, and meta-analyses were performed for subgroups with only two studies (min 11 patients), and provided subgroup data for only one eligible study in table 2. On the other hand, as many as 5 studies reported data for “cagPAI and oipA” as mentioned in the results section (last paragraph). The authors should either report the detail methods and results for the outcome (oipA and virulence factors), or remove them from this report and mention it will be studied and reported in another systematic review. We cannot just list some
studies without a proper method section and omit the results and analyses based on “may be limited sample size”. This is an example for selective reporting/reporting bias in meta-analysis.

Minor comments (Minor essential revisions):

1. Some English errors are noted and the report requires English editorial support.

2. On page 6, the statement “no comprehensive meta-analysis has...”. I assume the authors meant no meta-analysis has been published, unless the authors want to say that the meta-analysis was available but was not comprehensive.

3. In figure 1, 168 records were excluded after reviewing the titles and abstracts. Therefore, “123 excluded after title and abstract review” should be “123 irrelevant studies”.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests