Author's response to reviews

Title: Case-only study of interactions between metabolic enzymes and smoking in colorectal cancer

Authors:

Chunhong Fan (happydou_8@hotmail.com)
Mingjuan Jin (jinmj@zju.edu.cn)
Kun Chen (ck@zju.edu.cn)
Yongjing Zhang (mayun_zyi@163.com)
Shuangshuang Zhang (qtss124@sina.com)
Bing Liu (Liubing7362@sohu.com)

Version: 3 Date: 15 April 2007

Author's response to reviews: see over
April 13, 2007

Dear Editor,

Thank you very much for your response information. We are pleased to answer the questions of the reviewers and the manuscript (Ms No.: 1390729009131154) has also been extensively revised according to the comments (resubmitted online). Please contact me if you have any problems or questions regarding our manuscript, and hope that the MS is now acceptable.

Best regards,

Sincerely yours,

Chunhong Fan
Department of Epidemiology and Health Statistics,
School of Public Health,
Zhejiang University
Email: happydou_8@hotmail.com

PS. The answers to the comments are provided by 8 separated sheets.
Reviewer: Dr. Ladero

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Question #1: Although this is a case-only study, it would be interesting that the authors include some data on the distribution of the studied polymorphisms in the general Han population. The percentage of null GSTM1 genotypes is quite high (58.8%) and, according to the Hardy-Weinberg equilibrium, the supposed allele frequency of the CYP2E1 T allele seems also unexpectedly elevated, indeed for an Oriental population. The percentage of null GSTM1 genotype and the CYP2E1 T allele are quite high.

Answer: The focus of this study was to reveal gene-gene and gene-environment interaction with a case-only design. So the frequencies of distribution in genetic polymorphisms were obtained from case population. In the control, the frequency of null GSTM1 genotype was 54.5%, which is close to what was reported in Guangdong (54.3%) [1] and Uygur (53.2%) [2]. The frequency of CYP2E1 T allele was 24.31%, which is similar to what Gao et al. (22.91%) [3] and Wang et al. (25.69%) [4] reported.

Question #2: It should be easier and more informative to express tobacco use as pack-years, reflecting the life-dose of tobacco more than the heaviness or the duration of the smoking habit. Moreover, in this way, the authors could express the tobacco exposure as an unique parameter.

Answer: We take Dr. Ladero’s suggestion to use pack-years to express tobacco exposure. The results (the fourth paragraph on page 7 and the second paragraph on page 8), discussion (the second paragraph on page 11) and table 4 of the manuscript have been revised accordingly. (With red font)

Question #3: The dichotomous division in ethanol drinkers/non-drinkers is too rigid. It seems more adequate to use the daily dose of ethanol, thus classifying patients as non/moderate drinkers (i.e., less than 50 g ethanol a day) or heavy
Answer: In the revised version of manuscript, we reclassified patients as non/moderate drinkers (less than 50g ethanol a day) and heavy drinkers (more than 50g ethanol a day) according to the comments (see the fourth paragraph on page 7 and Table 2 in the revised version). (With red font)

Question #4: The patients were survivors of the studied cohort diagnosed with colorectal cancer. It should be noted that some dead individuals could have died of colorectal cancer, perhaps introducing a bias, as these subjects could have suffered from more aggressive forms of the disease.
Answer: Participants in this study were selected from a CRC cohort-study population. At baseline, all cohort members were invited to a screen by RPHA-FOBT and asked to complete a questionnaire. No participant refused to be interviewed except those who were out of town at that time. In total, 64,693 people responded to the questionnaires among a population of 75,842. Follow-up for incident cancer in the entire cohort was established by the cancer registration system and CRC report system in Jiashan County. 207 incident cases were recruited in the present study, and when we analyzed the exposure factors between those cases and deceased, there was no significant difference between them. So, our results are unlikely to be attributed to selection bias. We also discussed this issue in the second paragraph of our manuscript on page 12. (With red font)

Question #5: In page 10, first paragraph, the authors state that SULT1A1 and CYP1B1 pathways “may not be of importance in the metabolic activation of these carcinogens in vivo”. Is this the statement that the authors want to express?
Answer: This is only our conjecture about the outcome of the findings. It needs to be confirmed in a future study.
Minor Essential Revisions

Question #1: The feet for table 2 contains a mistake, as it is quite difficult to smoke alcohol... In some countries it may be inhaled after boiling any spirit, but I think that this is not the case in China.
Answer: The “alcohol smoking” was corrected to “alcohol consumption” in table 2 of manuscript. (With red font)

Question #2: Not being a native English speaker, I can only recommend a language revision before publishing.
Answer: Language revision has been done under Prof. Levine’s help, a native English speaker from America.

Reviewer: Margreet Lüchtenborg

Major Compulsory Revisions

Question #1: It would be helpful if the authors could elucidate the study design in more detail. It was unclear to me whether all subjects were interviewed or only the subjects diagnosed with primary (?) colorectal cancer that were included in this study. The latter may have implications for the possibility of bias with regard to the information obtained for cigarette smoking and alcohol consumption and the authors would need to address this. Also, more detail with respect to the applied questionnaire and what period of exposure it referred to would provide more insight.
Answer: The study design and the status of exposure have been elucidated in more detail in the first paragraph of manuscript on page 5. (With red font)

Question #2: Although, as the authors explain, a case-only design provides an efficient way to test for interactions, no data with regard to power are included. Considering the study is limited to 207 subjects, what is the likelihood of missing any interactions that may be present?
Answer: Software QUANTO was used to compute power of gene-gene and
gene-environment interaction in the present study [5]. Assumed α was 0.05, and all p-value was two-side. The prevalence of the CRC was 0.0002415, and the exposure rate for cigarette smoking was 0.382. The power was shown in table 1 and 2. Given the number of CRC cases (n=207), the case-only may detect modest interactive effects among genotypes and cigarette smoking.

Question #3: What do the authors think the observed interaction between SULT1A1 and CYP1B1 may mean? In particular, with respect to offering an explanation to this and to the smoking and CYP1B1 interaction on page 11, the authors mention hypotheses derived from in vitro studies, but fail to explain these hypotheses in the light of an interaction between genotypes or to reference these.
Answer: On page 9 and 11 of the revised manuscript, we have mentioned these. (With red font in revised MS)

Question #4: The authors suggest that colorectal cancer may develop in an estrogen-dependent manner. However, they do not report results from an analysis stratified by sex.
Answer: In the present study, 105 participants were male and 102 were female. The power for these subjects to detect a gene-gene interaction by case-only is about 76%, which is lower than 80%. So the results from stratified analysis by sex are inconsistent. Furthermore, these preliminary exploratory results should be confirmed in a larger study in future.

Minor Essential Revisions
Question #1: Although there is mounting evidence that smoking may be associated with colorectal cancer, recent expert reports (IARC, US Surgeon General reports on tobacco smoking and cancer) have not included colorectal cancer in the list of tobacco related cancers, because of inconsistencies
between studies and the potential for confounding by alcohol. Although the authors remark the absence of a clear causal relationship in the Discussion, the authors should rephrase the wording throughout the manuscript, so as not to overstate current evidence.

Answer: The relationship between smoking and CRC has rephrased, mainly in the first paragraph on page 3. (With red font in revised MS)

Question #2: On page 11, the authors write “Evidence suggests that … risk of CRC.” I think they omitted a reference.

Answer: According to another reviewer’s suggestions, pack-years was used to express tobacco exposure. So, discussion (on page 11) of the manuscript has been revised. “Evidence suggests that … risk of CRC” has also been deleted.

Question #3: The case-control study referenced in the second paragraph on page 12 was carried out by Slattery et al. in Northern California and Utah in the United States and not the Netherlands. A cohort study on smoking with a nested case-only study that found a borderline significant interaction between GSTM1 null and cigarette smoking was carried out in the Netherlands (Lüchtenborg et al., AJE, May 2005), and I leave it to the discretion of the authors to include this reference in the manuscript.

Answer: The mistake has been corrected accordingly, and the cohort study on smoking with a nested case-only study was referenced also. (With red font in revised MS)

**Discretionary Revisions**

Question #2: The authors correctly addressed the underlying gene-environment and gene-gene independence assumption for a case-only design. However, they may want to include in the Methods and Results section the testing of this assumption, of which the findings are reported in the Discussion.

Answer: According to another reviewer’s suggestions, we have added this in
the first paragraph of revised manuscript on page 5. (With red font)
Reference:


**qA:** the prevalence of the high-risk allele

### TABLE 1. Power for 207 subjects to detect a gene-environment interaction by case-only design

<table>
<thead>
<tr>
<th>qA</th>
<th>RgE</th>
<th>Rg1</th>
<th>Rg2</th>
<th>Rg3</th>
<th>Rg4</th>
<th>Rg1</th>
<th>Rg2</th>
<th>Rg3</th>
<th>Rg4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>2.0</td>
<td>0.38</td>
<td>0.53</td>
<td>0.54</td>
<td>0.52</td>
<td>0.06</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>3.0</td>
<td>0.80</td>
<td>0.91</td>
<td>0.89</td>
<td>0.86</td>
<td>0.08</td>
<td>0.11</td>
<td>0.13</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>0.96</td>
<td>0.98</td>
<td>0.98</td>
<td>0.95</td>
<td>0.10</td>
<td>0.16</td>
<td>0.20</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>0.99</td>
<td>0.997</td>
<td>0.99</td>
<td>0.98</td>
<td>0.14</td>
<td>0.22</td>
<td>0.27</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>2.0</td>
<td>0.69</td>
<td>0.58</td>
<td>0.44</td>
<td>0.33</td>
<td>0.28</td>
<td>0.42</td>
<td>0.47</td>
<td>0.48</td>
</tr>
<tr>
<td>3.0</td>
<td>0.97</td>
<td>0.87</td>
<td>0.70</td>
<td>0.54</td>
<td>0.66</td>
<td>0.83</td>
<td>0.85</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>0.99</td>
<td>0.95</td>
<td>0.80</td>
<td>0.62</td>
<td>0.88</td>
<td>0.97</td>
<td>0.97</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>0.999</td>
<td>0.97</td>
<td>0.84</td>
<td>0.66</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>2.0</td>
<td>0.49</td>
<td>0.29</td>
<td>0.19</td>
<td>0.14</td>
<td>0.63</td>
<td>0.66</td>
<td>0.57</td>
<td>0.48</td>
</tr>
<tr>
<td>3.0</td>
<td>0.81</td>
<td>0.49</td>
<td>0.31</td>
<td>0.21</td>
<td>0.96</td>
<td>0.95</td>
<td>0.87</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>0.91</td>
<td>0.58</td>
<td>0.36</td>
<td>0.24</td>
<td>0.997</td>
<td>0.99</td>
<td>0.95</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>0.95</td>
<td>0.63</td>
<td>0.38</td>
<td>0.25</td>
<td>0.999</td>
<td>0.996</td>
<td>0.97</td>
<td>0.89</td>
<td></td>
</tr>
</tbody>
</table>

**D:** Dominant; **R:** recessive

### TABLE 2. Power for 207 subjects to detect a gene-gene interaction by case-only design

| qA | qH | G | H | Rg=1 | Rg=2 | Rg=3 | Rg=1 | Rg=2 | Rg=3 | Rg=1 | Rg=2 | Rg=3 |
|----|----|---|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0.05 | 0.05 | D | D | 0.19 | 0.40 | 0.55 | 0.48 | 0.84 | 0.94 | 0.74 | 0.98 | 0.995 |
| 0.05 | 0.25 | D | D | 0.39 | 0.51 | 0.51 | 0.81 | 0.89 | 0.86 | 0.96 | 0.98 | 0.96 |
| 0.05 | 0.25 | R | R | 0.14 | 0.31 | 0.46 | 0.35 | 0.72 | 0.89 | 0.58 | 0.93 | 0.99 |
| 0.25 | 0.25 | D | D | 0.69 | 0.55 | 0.40 | 0.96 | 0.84 | 0.64 | 0.995 | 0.92 | 0.73 |
| 0.25 | 0.25 | R | R | 0.28 | 0.40 | 0.43 | 0.66 | 0.81 | 0.81 | 0.89 | 0.95 | 0.94 |