Author's response to reviews

Title: Analysis of EGFR, HER2, and TOP2A gene status and chromosomal polysomy in gastric adenocarcinoma from Chinese patients

Authors:

Zhiyong Liang (liangzhiyong1220@yahoo.com)
Xuan Zeng (zengx88@yahoo.com.cn)
Jie Gao (zy_795@hotmail.com)
Shafei Wu (Suphie@sina.com)
Peng Wang (chinaroc@tom.com)
Xiaohua Shi (shixiaohua3762@hotmail.com)
Jing Zhang (jingzhangsmile@163.com)
Tonghua Liu (liuth_pumch@yahoo.com.cn)

Version: 2 Date: 24 June 2008

Author's response to reviews: see over
Dear editors,

Thank you very much for your email about our manuscript (MS: 4555040261914632), we have carefully read the suggestions from you and reviewers, and try to answer the questions point-by-point.

Answers to editor:

1. The title of our manuscript has been changed to “Analysis of EGFR, HER2, and TOP2A gene status and chromosomal polysomy in gastric adenocarcinoma from Chinese patients”

2. We have used a commercial copyediting service to improve our English, abstract has been revised.

3. The terminology and controls we use in our manuscript is consistent with the reporting of our data.

4. Authors' contributions and Acknowledgements sections have been added into our revised manuscript.

Answers to reviewer 1:

**Question 1.** In figure 1, 2 and 3 negative/underexpressed controls for immunohistochemistry need to be included.

**Answer:** Negative/underexpressed controls for immunohistochemistry staining of EGFR, TOP2A and HER2 have been added to figure 1, 2, and 3.

**Question 2.** The numbers in text on the results section of page 10 section
3.2 do not match with table2.

Answer: We are very sorry for our mistake about the numbers in text of section 3.2 did not match with table2. We have carefully checked our original data and revised table2 with correct number.

Question 3 .. The manuscript has lot of syntax and spelling errors The following minor corrections need to be included

（1）Abstract Section

On Page2 Line 4 The sentence

“No detailed report about the above three genes in chinese patients, so we try to determine the gene status of EGFR, Her2 ,Top2A genes and anuesomy of chromosome 7 , 17 in chinese gastric carcinoma patient”. Should be rephrased On Page 2 Line 8 the sentence can be written as Gastric carcinoma patients were investigated by using immunohistochemistry Gastric carcinoma patient tissue samples were investigated by immunohistochemistry On Page 2 Line 21 Conclusion,Remove comma after conclusion .

Answer: Abstract section has been revised according to reviewer’s suggestion.

“No detailed report about the above three genes in chinese patients, so we try to determine the gene status of EGFR, Her2 ,Top2A genes and anuesomy of chromosome 7 , 17 in chinese gastric carcinoma patient” has
been changed to “Herein, we investigated the gene status of EGFR, HER2, and TOP2A in Chinese gastric carcinoma patients. We determined the rate of polysomy for chromosomes 7 and 17, and we attempted to clarify the relationship between EGFR, HER2, and TOP2A gene copy number and increased expression of their encoded proteins. Furthermore, we tried to address the relationship between alterations in EGFR, HER2, and TOP2A and chromosomal polysomy.”

On Page 2 Line 8 the sentence “Gastric carcinoma patients were investigated by using immunohistochemistry” has been changed to “Gastric carcinoma patient tissue samples were investigated by immunohistochemistry”

On Page 2 Line 21 Conclusion, comma was removed after conclusion.

( 2 ) **In Introduction section**

On Page 3 line 7

Recently target therapy has been applied -should be changed to

-Recently targeted therapy has been applied

On Page 3 line 8

Current target -should be changed to

-Current targeted

**Answer :** Introduction section has been revised according to reviewer’s suggestion. On Page 3 line 7, “Recently target therapy has been applied”
changed to “Recently targeted therapy has been applied”

On Page 3 line 8, “Current target” has been changed to “Current targeted”

(3) On Page 28 Legend to figures section

Figure 2 B: Immunohistochemistry showed Top2A NI (nuclear index)

Immunohistochemistry showed Top2A NI (nuclear index) In figure 4 the heading TOP2A Expression Nuclear Index (NI) has to be changed to TOP2A Expression Nuclear Index (NI)

Answer: Page 28 Legend to figures section, Figure 2 B:

“Immunohistochemistry showed Top2A NI (nuclear index) “has been changed to “Immunohistochemistry showed Top2A NI (nuclear index)” In figure 4 the heading “TOP2A Expression Nuclear Index (NI)” has to be changed to “TOP2A Expression Nuclear Index (NI)”.

Answers to reviewer 2:

Major Compulsory revisions

Question 1. The word "polysomy" is defined as an excess of a particular chromosome. Therefore, the terms "EGFR polysomy" and "EGFR aneusomy (non-disomy)" make no sense to me. Please carefully define and explain the meaning of these terms.

Question 2. p8: The EGFR gene status was classified into six categories: disomy, low trisomy, high trisomy, low polysomy, high
polysomy and gene amplification. In this categorization, two standards are used. Polysomy 7 and the EGFR gene amplification are not mutually exclusive phenomena. For example, cases with either "gene amplification and disomy" or "gene amplification and polysomy" could also exist. Of course, this confusion could be avoided by changing the present categorization. In general sense trisomy is a low polysomy. The reason why trisomy was isolated from low polysomy is unclear and difficult to understand. Furthermore, it is extremely unclear regarding the reason for dividing trisomy into two categories, namely "low" and "high". Tables 1 and 2 have no column of "high trisomy", in these tables only 5 categories are observed. Furthermore the discussion does not mention anything about high and low trisomies.

**Answer : We try to answer question 1 and 2 together.** We agree that the word “polysomy” was originally defined as an excess of a particular chromosome. “Chromosome polysomy” and “EGFR polysomy” is different in our study, “Chromosome polysomy” refers to high chromosome copy number, “EGFR polsyomy” refers to high EGFR copy number. Because there is no standard interpretation of EGFR gene status by FISH analysis, so “polysomy” was introduced into interpretation of EGFR gene status in several studies, the widely accepted score and category system was established by Cappuzzo F et al.
Growth Factor Receptor Gene and Protein and Gefitinib Sensitivity in Non–Small-Cell Lung Cancer. Journal of the National Cancer Institute, 2005, 97(9):643-655. Patients were classified into six FISH strata with ascending number of copies of the EGFR gene per cell according to the frequency of tumor cells with specific number of copies of the EGFR gene: 1) disomy (≤ 2 copies in >90% of cells); 2) low trisomy (≤ 2 copies in ≥ 40% of cells, 3 copies in 10% – 40% of the cells, ≥ 4 copies in <10% of cells); 3) high trisomy (≤ 2 copies in ≥ 40% of cells, 3 copies in ≥ 40% of cells, ≥ 4 copies in <10% of cells); 4) low polysomy (≥ 4 copies in 10% – 40% of cells); 5) high polysomy (≥ 4 copies in ≥ 40% of cells); and 6) gene amplification (defined by presence of tight EGFR gene clusters and a ratio of EGFR gene to chromosome of ≥ 2 or ≥ 15 copies of EGFR per cell in ≥ 10% of analyzed cells).

For further analysis, the above six groups were reclassified into two categories by Cappuzzo F et al.: FISH positive(high polysomy and amplification), FISH negative(disomy, low trisomy, high trisomy, low polysomy).

According to the criteria of Cappuzzo F et al, EGFR FISH positive results were not consistent with EGFR protein overexpression in previous and our studies. In our present study, we tried to determine whether the alteration of EGFR copy number is related to EGFR expression by FISH analysis, we statistically analyzed different combination of the above six
groups according to the criteria of Cappuzzo F et al, our FISH results showed that non-disomy combination (include low trisomy, high trisomy, low polysomy, high polysomy and amplification groups) was significantly consistent with the overexpression of EGFR protein, this is the first report about the relationship between EGFR non-disomy (aneusomy) and EGFR overexpression, and it need to be confirmed by further studies.

Because no high trisomy was detected in our results, so we omitted the “high trisomy” column in table 1 and 2, we have added the “high trisomy” column in table 1 in our revised manuscript. Table 2 has been revised, high trisomy column is not needed

**Question 3.** The authors also used an additional categorization #EGFR-positive and -negative. The authors stated in the abstract (p2, line 17) "Chromosome 7 polysomy correlated significantly with EGFR FISH-positivity." This line of reasoning is very difficult to understand because EGFR FISH-positive includes high polysomy by definition. These complicated and duplicated categorizations thus lead to a great deal of confusion.

**Answer :** According to the criteria of Cappuzzo F et al, “EGFR polysomy” was different to “Chromosome 7 polysomy”, EGFR FISH-positive category include “EGFR high polysomy and amplification”, these categorizations have been widely used in the studies
about EGFR analysis.

Our result showed that "Chromosome 7 polysomy correlated significantly with EGFR FISH-positivity.", this indicate that chromosome 7 polysomy may be one of the reasons which caused high copy number of EGFR gene.

**Question 4.** One of the major findings of this paper is that not the EGFR FISH-positive status but EGFR aneusomy showed a significant association with EGFR overexpression (p10 & p13). Does this means that trisomy 7 and low-polysomy 7 have a major impact on the EGFR overexpresion ? Authors should clearly state which of the 6 categories for the EGFR gene status is supposed to be closely associated with EGFR overexpression.

**Answer**: We have statistically analyzed different combinations of the 6 groups for the EGFR gene status according to the criteria of Cappuzzo F et al. In our results, only the non-disomy combination (included low trisomy, high trisomy, low polysomy, high polysomy and amplification) showed significant association with EGFR overexpression, the "non-disomy combination” was classified as “EGFR aneusomy “in our study, the “EGFR aneusomy “ was different to “chromosome 7 aneusomy”, the result indicated that EGFR trisomy, polysomy and amplification may all have impact on the EGFR overexpression. This need to be confirmed by further studies.
Minor Essential Revision

**Question**. Reference 2: second author "Kunomo K" --> "Kunitomo K"

**Answer**: Reference 2: second author "Kunomo K" has been changed to "Kunitomo K".

Discretionary Revision

**Question**. On page 14, line 16: "a small number of cases showed co-expression of EGFR and HER2." In this section, the precise number of cases should be written. Thereafter, an expanded discussion of the dual inhibition therapy against EGFR and HER2 would be possible.

**Answer**: "only a small number of cases showed co-expression of EGFR and HER2, and a small number of cells were simultaneously FISH-positive for EGFR and HER2." has been changed to "only 7 cases showed co-expression of EGFR and HER2, and 3 cases were simultaneously FISH-positive for EGFR and HER2. Dual inhibitor of EGFR and HER2 may be benefit to patients with gastric carcinoma."

Answers to reviewer 3:

Major Compulsory Revisions

**Question 1.** It is not clear what a Spearman test is involved. Most of the comparisons performed in the report are between two binary measures (e.g., overexpression versus not, disomy versus aneusomy), which can be
done by Pearson’s Chi-square test or Fisher’s exact test (when small cell count exists). Meanwhile, it is also useful to provide a measure of association besides the results from a significance test. No spearman correlations or other measure of associations were reported. Because all the measures are often binary, or only have a few categories (<5), other measure of associations such as odds ratio would be more is appropriate than a spearman correlation.

**Answer:** According to reviewer’s suggestion, all the comparisons have been statistically analyzed by Pearson’s Chi-square test or Fisher’s exact test, and odds ratios have been calculated.

**Question 2.** There are a lot of statistical comparisons done in this report. The issue of multiple comparisons should be noted and adjusted.

**Answer:** The comparisons have been adjusted, and each comparison has been given a notice in revision.

**Question 3.** It is not clear why they wish to establish cut point(s) for NI and why only 5%, 10%, and 25% were explored. Can a continuous NI be also informative? A better and common method to establish cut points for these data would be using Receiver Operating Characteristic (ROC) curve analysis. Table 5 needs a better format and less information.

**Answer:** The relationship between TOP2A protein expression and TOP2A gene amplification was not clarified in previous studies, the purpose of setting three cut points for TOP2A NI in our present study was
to determine which one would be consistent with TOP2A amplification.

Based on previous reports (Hansel DE et al. a subset of pancreatic adenocarcinomas demonstrates coamplification of topoisomerase IIα and HER2/neu. Am J Clin Pathol 2005,123:28-35), we selected 5%, 10% and 25% as cut off points, and tried to identify the difference in the presence of TOP2A amplification with cutoff of less than 5%(10%, 25%) vs 5%(10%, 25%) or more. Table 5 has been modified and renamed as Table 6, while Table 6 has been renamed as Table 5 in revision.

**Question 4.** Figure 4 can be better presented with a histogram, which is able to show many characteristics of a distribution (e.g., central location, spread, and range). Those descriptive statistics should be reported. The bin for the histogram should be properly selected to capture the variability of the distribution.

**Answer:** We used Figure 4 to present the TOP2A nuclear index of 100 cases of gastric carcinoma, and added arrows to show the 3 cases with TOP2A amplification. We also revised the legend.

**- Minor Essential Revisions**

The reporting of the results is generally difficult to follow and sometimes confusing. Both the texts and tables for results require improvement.

**Answer:** Both the texts and tables for results have been revised.

**Question 1.** It is not clear by the first sentence in the statistical analysis
section when they wish to compare ECFR overexpression what do they consider as positive and negative tumors.

**Answer:** Immunohistochemical staining for EGFR was evaluated following the criteria recommended by the manufacturer: 0, no discernible staining or background type staining; 1+, equivocal discontinuous membrane staining; 2+, unequivocal membrane staining with moderate intensity; and 3+, strong and complete plasma membrane staining. More than 10% of the cells were required to meet the criteria for EGFR analysis. Scores of 2+ and 3+ staining levels were considered to be EGFR overexpression. These has been introduced in “Materials and Method” section.

**Question 2.** Table 1, 2, 4, 5: It is confusing that the table lists all the categories but p-value were come from the statistical tests based on combined categories (e.g., 0, 1+,2+, 3+, versus overexpression (2+, 3+) versus not (0, 1+)). And different ways of combining categories gave different p-values. Alternative presentation of the table/results should be explored. For example, one can present the combined categories with associated p-values in a table, or present all categories in table, but in the text, present p-values along with the N or percentage of the combined categories.

**Answer:** All the tables and texts in Results section have been revised according to reviewer’s suggestion.
Question 3. Footnotes for Table 1, 2, 4: the word “difference” is not specific and could be misleading. Depending what statistical tests are used, one can replace it as “association”, “independence”, or “correlation”. The two footnotes for Table 2 are duplicated.

Answer:  Footnotes for Table1,2,4 has been revised.

Question 4. Table 3, the p-value would never be 0. Replace 0.0000 with a small number or express it as <0.001 or <0.0001.

Answer:  the p-value 0.0000 in table 3 was replaced as p<0.001

Question 5. More explanation on samples and the sampling process is needed. How many patients were eligible to be selected for the period of 2000-2005 for this study? How representative of the patient population seen at the Department of Surgery in the Peking Union Medical College Hospital as compare to the Chinese population that seek their treatment in other parts of country? Is there evidence that the random selection has produced a sample as expected?

Answer:  793 patients were eligible to be selected for the period of 2000-2005 for this study. More than half of the patients seen at the Department of Surgery in the Peking Union Medical College Hospital were not Beijing citizens, they came from different regions of China. We selected the specimens from our archive paraaffin embedded blocks. All the specimens were selected by two pathologists and only those patients whose clinical data intact and the blocks were enough
to be cut into 20 slides were selected.

**Question 6.** How the sample size of 100 was determined? What is the minimum association (e.g., odds ratio, correlation) can be detected with 100 subjects for 80% or 90% power? This information can inform the readers that a small or significant p-value is not simply a consequence of a large sample size.

**Answer:** The sample size of 100 was initially determined by Dr. Tonghua Liu and Dr Zhiyong Liang before the study. We are very sorry that no minimum association was calculated.

**- Discretionary Revisions**

**Question 1.** For all tables, it should be simple and clear enough to use “Total” in replace of “Total number”. In addition to column total, row total can be listed as well.

**Answer:** “Total number” has been changed to “Total” in all tables, and row totals were listed in some tables.

**Question 2.** Move up the result for the correlation with clinicopathologic variables, section 3.9, as this was the first analysis the authors proposed to perform.

**Answer:** Section 3.9 in results has been moved up. The sequence number for other sections were all revised.