Peer Review File

Article Information: https://dx.doi.org/10.21037/jgo-23-735

Reviewer A

This is a study using a novel in-house method to capture CTC in CRC patients as a liquid biopsy using retrospective data. Whilst this is an interesting premise, I have several concerns with this manuscript.

1. *Overall needs language editing and re-writing. It is not well written - multiple grammatical, punctuation and editing errors. Needs clarification of acronyms used*

Reply 1: Thank you for your advice. Based on your suggestions, we have touched up the content of the article, reworked the grammar, punctuation, and editing errors, and standardized the acronyms used.

2. *Whilst Cellsearch does have multiple limitations, there are several other methods of CTC capture with improved sensitivity and specificity (ref https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8716996) - why did the authors choose their method?*

Reply 2: Despite Cellsearch's limitations, a large number of clinical trials continue to use the CellSearch platform, which reflects the robustness of the Cellsearch platform and the positive enrichment it represents (***Sotelo, M. J. et al. Ann Oncol 26, 535–541 (2015)***). In addition, the CellSearch platform has difficulty in detecting CTCs in the EMT state as well as those of mesenchymal origin, and many studies have suggested that the capture of CTCs of these origins could be enhanced by the addition of mesenchymal antibody-modified magnetic beads (***Yi, B. et al. J Nanobiotechnology 19, 74 (2021); Cha, J. et al. Cancers (Basel) 15, 2825 (2023).***). Since the customization of magnetic beads is relatively flexible and free, we can use specialized antibody-modified magnetic beads to conduct in-depth studies on colorectal cancer phenotypes of interest.

3. *Materials and methods*

- Unclear inclusion and exclusion criteria for patient selection.

- It is suddenly stated in the results section that 25 non CRC patients were also included!

- Difficult to follow authors' in-house method of CTC capture

Reply 3: As per your request, we have added inclusion and exclusion criteria. (see Page 5, line 99-104). And based on your suggestion, we have added the recruitment of healthy volunteers in the methodology section (see Page 5, line 106-108). In order to more explicitly describe our CTC capture methods, we have added relevant content (see Page 7, line152-158).

4. *Results + discussion*

- Stated sensitivity fluctuates in the manuscript from 90% to 76%. Which is it?

- The stated endpoint in patient OS - however this data is never presented

- There was no association between CTC and any clinicopathological parameter (Table 2). Why was this the case? And how then were there suddenly significant findings with CTC as a

independent prognostic factor on regression analysis? This needs explanation and review of the data/statistics

- what is the clinical importance of this paper? We already now that CTCs have significance in cancer prognosis - this paper does not add to this.

Reply 4: We carefully examined the full text and confirmed that the sensitivity of the system is 79.0%, and corrected the typographical errors (see Page 1, line 29-30). Based on your comments, we added data on patient OS (see Page 9, line 228-229). We have noted the statistical results on CTC in TABLE 2, and we have re-examined the statistics and we suspect that this phenomenon occurred because we included the total number of mixed epithelial CTCs and mesenchymal CTCs in our statistical calculations. CTCs have significant heterogeneity, and it is currently possible to categorize CTCs into epithelial and mesenchymal types based on the antibodies that are expressed by the CTCs, and studies have found that both types of CTCs have significant heterogeneity during the metastatic process. The phenotypes of the two types of CTCs in the metastatic process are very different, epithelial CTCs are associated with distal metastasis formation, however, mesenchymal CTCs are less able to form metastatic foci, and they are mainly associated with chemoresistance of the tumor. Therefore when mixing two CTCs with different invasive abilities some of the characteristics tend to be masked. We re-analyzed epithelial CTCs alone with the baseline characteristics of the cohort and found that CTCs were significantly associated with the pathological stage of AJCC in the cohort. We added relevant content to the Discussion section and provide a table of the relationship between epithelial CTC and clinicopathologic features inside the Supplementary File (see Page 11, line 271-287 and Page 28 line 480). Our study also showed that a CTC classifier integrating non-invasive and reproducible features is more feasible and less costly than prognostic features in previous studies. Unlike traditional nomograms that use clinical prognostic factors, since CTC levels can reflect the malignancy of the primary tumor we also incorporated CTC counts into a nomogram that predicts the survival risk of individual patients. Our statistical model demonstrated that elevated CTC counts were shown to be significantly associated with lower survival in patients with CRC, so incorporating CTC counts into this nomogram helped to improve its predictive accuracy. This will pave the way for the creation of a simple and accurate prognostic prediction method for CRC patients. we have add our text as advised (see Page 12, line 297-305).

Reviewer B

In this manuscript, the authors investigated the significance of evaluating circulating tumor cells (CTCs) in colorectal cancer (CRC) patients using in vitro and clinical data. As a result, CTCs were recognized as a novel biomarker and reported to be an independent prognostic factor, potentially assisting in the creation of a monogram.

However, the current manuscript is not yet sufficient for publication in the "Journal of Gastrointestinal Oncology." The authors will need to revise this manuscript in accordance with the comments provided.

1. *Clinical-pathological factors being presented are insufficient. While there are many Stage IV cases, it is unclear whether R0 resection was achieved, the location and number of*

distant metastases, the presence of RAF/BRAF mutations, TNM factors, and the use of neoadjuvant chemotherapy. At least Tables 1 and 2 should include this information. Since CEA is included in Tables S1 and S2, CEA should also be included.

Reply 1: Thank you for your comments. In Tables 1 and 2 we have added the clinicopathologic factor information as advised (see Page 18-21, line 441-445).

2. *In Table 2, what is the basis for setting the CTCs cutoff at 5? Based on Table 1, it seems necessary to evaluate with a median of 7. Additionally, is it confirmed that these CTCs are the number collected with EP+Vi-IMB?*

Reply 2 : To assess the ability of the CTC assay to discriminate between patients and healthy volunteers, ROC curves were used. The area under the ROC curve was 0.872 (Figure 3E). We defined the threshold as 5 CTCs per 7.5 ml of blood using youden index. We have add our text (see Page 9, line 220-221). We have carefully checked the data and confirmed that these CTCs are the number collected with EP+Vi-IMB.

3. To assess Total CTC as a prognostic factor, please create Kaplan-Meier curves.

Reply 3: We have added the Kaplan-Meier curves as advised (see Page 27, line 472).

4. *The subjects are patients from 2016-2019, but the median follow-up is only 2 years, which is a short dataset. What is the reason for this? Is it because many Stage IV patients died of cancer, or were there dropouts?*

Reply 4: We reviewed the survival data of the cohort, and as you said, most of the stage IV patients in the cohort died of cancer, and we will follow up by further expanding the number of cases enrolled to improve the accuracy of the median follow-up time and median overall survival

5. *Perform the experiment in Fig. 2G with cell lines other than LOVO and consider recommended dosages.*

Reply 5: Following your suggestion, we repeated the experiment with HT29, reanalyzed and plotted the results based on the experimental data, and showed that the optimal dosage for the preparation of the immunomagnetic beads was still 60 μg. we have modified our text and Figure as advised (see Page 8, line 197-198 and Page 23, line 449)

6. For the comparison of ROC curves in Fig. 3E and F, it would be better to compare EP-IMB or Vi-IMB individually, showing the superiority of EP/Vi-IMB, rather than using CEA.

Reply 6: We have modified relevant pictures and content as advised (see Page 9, line 216-220 and Page 25, line 458)

7. *In the main text, it is stated that EP-IMB/Vi-IMB has the highest cell recovery rate (line 166), but based on Fig. 2C and D, it seems that EP+Vi-IMB has the highest cell recovery rate. Could you clarify this?*

Reply 7: We have carefully examined the part of the article you mentioned. We found that line 166 of the article mentions that " The most efficient capture was achieved when Ep-IMB/Vi-IMB was added sequentially with …". It actually refers to EP-IMB+VI-IMB, but this way of writing can easily be confused with EP-IMB/VI-IMB, so we have modified the description to "The Ep-IMB+Vi-IMB group has the highest capture efficiency when the total number of beads is kept constant". we have modified our text as advised (see Page 8, line 188-189).

8. *In Table S1, please add explanations for the cutoff values of CEA, Age, and Total CTC.*

Reply 8: We have added explanations for CEA, Age and Total CTC cutoff in the revised version(see Page 28, line 477).

9. *In Table S2, when conducting multivariate analysis, there may be a strong correlation between the N factor and Stage. Is it valid to evaluate CTC in this multivariate analysis, or should it be confirmed by a statistician?*

Reply 9: When performing multivariate analysis we also noticed that there may be some correlation between N factor and stage, in order to construct a more accurate COX regression model, we gradually removed variables according to the Akaike information criterion (AIC), compared different models, and finally chose the one with the smallest criterion value, and the model finally chose these four variables to participate in the modeling.

10. *Please add error bars to Fig. 3D.*

Reply 10: We have added the error bars based on your suggestion (see Page 25, line 458).

11. *Minor language corrections are needed. Abbreviations such as EP-IMB, Vi-IMB are used without explanation. There is a typographical error in CA19-9 (line 59), and there is insufficient consistency in terminology in the tables (Sex and Gender).*

Reply 11: We have added explanations of the abbreviations (see Page 1, line24-25), corrected the typographical errors (see Page 3, line 52), and harmonized the terminology in the table (see Page 20, line 444).