Peer Review File

Article information: https://dx.doi.org/10.21037/tlcr-22-801

Reviewer A

The authors offer study of circulating tumor cells collected from different anatomical sites from patients with pre-surgical non-small cell lung cancer. The study is reasonably powered with 119 patients available for analysis. The authors forgo any enrichment steps other than enrichment of mononuclear cells which are then subjected to nested-PCR. They offer validation data that shows that this gene expression correlates with the number of cells spiked into blood although they do not test the limits of detection below about 10 cells. The manuscript is concise and reasonably well written.

The intro mentions there is only 1 FDA cleared method for CTC enumeration. While correct as written, it may be worth noting that the Angle Parsortix platform was recently FDA cleared for CTC enrichment (https://pubmed.ncbi.nlm.nih.gov/36358657/)[1].

Reply: Agreed. The Parsortix reference was added and cited[1].

See: page 5, lines 106-109.

Methods, The authors note that they follow many of the preferred practices such as discarding the first 3 mL of peripheral blood collected to minimize epithelial contamination; however the typical time from collection to processing is noted only as “immediately.”

Reply: Agreed. The time frame was specified. The *sentence “The average time from blood sampling to processing did not exceed two hours”* was added.

See: page 6, lines 140 – 141.

Are the cells subject to analysis mononuclear cells or total white cells? Mononuclear cells usually require a density gradient centrifugation step that is not listed in the methods.

Reply: The total white blood cells, containing polymorphonuclear and mononuclear cells as well, were used for the CTCs detection. No density gradient centrifugation was performed because of the risk of CTCs lost.

The manuscript was revised to be more clear and not confusing and the sentence “*The total white blood cells were isolated by osmotic lysis from peripheral blood and bone marrow samples”* was added.

See: page 8, lines 175-176

Please clarify method for spiking cells. Based on the data in figure 1, this appears to be a limiting dilution.

Reply: The paragraphs describing spiking experiment and its results was edited to be clearer. In principle we used direct sorting of H2228 cancer cells by flow cytometry into white blood cell aliquots, so it was not limiting dilution. Honestly, it is extremely difficult biologically, and technically, to reliably prepare experimental samples containing one live cancer cell in 10 million non-cancer cells. In our study, we reproducibly detected 10 cancer line cells in 10 million buffy coat cells both by CFA and qRT-PCR methods.

The qRT-PCR methods are well established and known for their extreme sensitivity achieving

the detection rate 1:10 000 000 (Andergassen U.,et al., Oncology Reports, 2016; Wang S., et al, PLoS One, 2014)[2,3] enabling reliable amplification and detection of one cancer specific molecule in more than 10 million of others. It has been further described the higher expression of many biomarkers including CEA and EpCAM in primary tumor cells in comparison with cancer cell lines. In Protein Atlas (www.proteinatlas.com), the normalized CEA transcript expression of lung cancer is 30.2 in comparison to 3.3 in H2228 cell line. Kim N., et al. described the 3-6x fold time higher CEA expression in lung tumor tissue in comparison with H2228 cell line (BMC Cancer, 2022)[4] .

Based on database and published results, we assume that qRT-PCR method for CEA and EpCAM in lung cancer is sensitive enough to detect CTCs at levels 1-10: 10 000 000.
See: page 10, lines 228-235; page 20, Figure 2 legend, lines 491-492; page 15, lines 382-383

Combination of EGFR and CEA shown to significantly impact OS. Is this similar to the multivariate analysis in line 366 that seems to be described in the methods (line 278)? Is it reasonable that EGFR has a HR significantly less than 1 in the adjusted overall survival analysis (HR = 0.34, table 4), but in combination with CEA the prognosis is worse (HR 1.88)?
Reply: It was our misunderstanding. The results of two different models (Adjusted and Multivariate Cox regression models) were combined in one sentence. The sentence *"Furthermore, the EGFR mRNA-positive CTCs in the tumor draining pulmonary blood was shown to significantly impact DFS and OS (HR=0.36, P=0.020; HR=0.34, P=0.010, respectively) in adjusted survival analysis"* was edited and moved at proper place.
The Multivariate Cox regression model identified only the presence of CEA mRNA-positive CTCs/DTCs in the peripheral blood and bone marrow at surgery as an independent negative prognostic factor for DFS.
See: page 14, lines 352-354

Abstract:
Please clarify the p-values for CSS and DFS, only one is listed here.
Reply: Both p-values were listed *"Patients with the presence of epithelial cellular adhesion molecule (EpCAM) mRNA-positive CTCs in TDB samples had significantly shorter CSS and disease-free survival (DFS) (p<0.031, resp. p<0.045)"*
See: page 3, lines 60-61

Grammar:
Line 250: cells were spiked into 10 million buffy coats should be "buffy coat <b> cells "
Reply: Agreed. The text has been edited.
See: page 10, lines 228, 232, 235 and 238

The manuscript is concise and reasonably well written but the brevity makes a few points difficult to grasp.

Reviewer B

This is a study which falls into the scope of the journal. The study has been conducted systematically with most of the areas covered in terms of the diagnostic and prognostic factors in NSCLC.
I have some concerns.

1. How sensitive is this method? Since only 10 H2228 cells are spiked in leukocyte preparation, can the RNA coming from these 10 cells enough to bring out the differences in epithelial cells/tumor-specific gene expression?
Reply: It is extremely difficult biologically, and technically, to reliably prepare experimental samples containing one live cancer cell in 10 million non-cancer cells. In our study we used direct sorting of H2228 cancer cells by flow cytometry into white blood cell aliquots and we reproducibly detected 10 cancer line cells in 10 million buffy coat cells both by CFA and qRT-PCR methods.
The qRT-PCR methods are well established and known for their extreme sensitivity achieving the detection rate 1:10 000 000 (Andergassen U.,et al., Oncology Reports, 2016; Wang S., et al, PLoS One, 2014)[2,3] enabling reliable amplification and detection of one cancer specific molecule in more than 10 millions of others. It has been further described the higher expression of many biomarkers including CEA and EpCAM in primary tumor cells in comparison with cancer cell lines (www.proteinatlas.com). Kim N., et al. described the 3-6x fold time higher CEA expression in lung tumor tissue in comparison with H2228 cell line (BMC Cancer, 2022)[4] .

Based on the database and published results, we assume that qRT-PCR method for CEA and EpCAM in lung cancer is enough sensitive to detect CTCs at levels 1-10: 10 000 000 providing thus more than sufficient sensitivity.
See: page 10, lines 228-235; page 20, Figure 2 legend, lines 491-492; page 15, lines 382-383

2. Among the study patients who died due to lung cancer, is the experimental and clinical data available (CTCs/DTC and other parameters)? This will presumably reveal the role of these factors in NSCLC-associated survival.
Reply: The comprehensive clinical, molecular and experimental data are provided in the Tables 1 and 2 and Figure 3, the most important data are further mentioned in the text (See Result). In our study, the presence of CTCs/DTCs did not correlate with other clinic-pathological parameters (page 13, lines 317-319). A statistically significant correlation was observed only in the case of smoking, but the patient population was unevenly distributed (page 12, line 319-322). The presence of a CTCs/DTCs has been identified as an independent prognostic factor in our study.

3. Has the follow-up been done in the study subjects? If yes, how many patients with higher number of CTCs/DTCs had recurrence. This would be an important factor in predicting the recurrence and also tailoring treatments directed towards micrometastatic disease responsible for NSCLC recurrence.
Reply: The median follow-up was 41.6 months in our study (Table 1; Page 12, line 307). The recurrence rate is displayed in Figure 4. For example, the 48.4% (15/31) patients with CTCs positivity in peripheral blood had an event (DFS).
See: Table 1; Figure 4

4. Although data has been provided extensively in the form of table, it might be useful to graph some of these data which makes it clear for the readers.
Reply: Agreed. Table 3 was converted into the box-plots (Figure 3).
See: Figure 3

References:

1. Cohen EN, Jayachandran G, Moore RG, et al. A Multi-Center Clinical Study to Harvest and Characterize Circulating Tumor Cells from Patients with Metastatic Breast Cancer Using the Parsortix® PC1 System. Cancers 2022 Oct 26;14(21):5238. doi: 10.3390/cancers14215238. PMID: 36358657; PMCID: PMC9656921.
2. Andergassen, U., Kölbl, A. C., Mahner, S., et al. Real-time RT-PCR systems for CTC detection from blood samples of breast cancer and gynaecological tumour patients. Oncology Reports 2016, 35 (4)1905-1915. https://doi.org/10.3892/or.2016.4608
3. Wang S, Zheng G, Cheng B, et al. Circulating Tumor Cells (CTCs) Detected by RT-PCR and Its Prognostic Role in Gastric Cancer: A Meta-Analysis of Published Literature. PLOS ONE 2014. 9(6): e99259. https://doi.org/10.1371/journal.pone.0099259
4. Kim N, Jeong D, Jo A, et al. Prescreening of tumor samples for tumor-centric transcriptome analyses of lung adenocarcinoma. BMC Cancer 2022 Nov 17;22(1):1186. doi: 10.1186/s12885-022-10317-9. PMID: 36397035; PMCID: PMC9673386.