Peer Review File

-Reviewer A-

We are pleased that Reviewer A is positive about our paper and we thank her/him for providing the below suggestions.

This is a very well written paper. The subject is not that novel but still of interest.

Similar results have been previously generated in other tumours including BRAF mutant melanomas treated with BRAF inhibitors showing no significant difference in overall survival (OS) where patients were divided to high and low allele frequency (AF) groups. The result on progression free survival (PFS) can be slightly variable from one study to another, and can be affected by the number of cases included, the accuracy of calculation of allele frequency in relation to tumour cell content, molecular methods used and the overall study design.

But it makes sense biologically that tumours with high AF of mutant EGFR showing a better response to EGFR inhibitors initially, hence better PFS, and eventually deteriorate due to expansion of non-EGRF mutant clones or clones acquiring further mutations downstream EGFR or in other pathways, hence similar OS compared to low AF tumours.

Comment #1: The main criticism/limitation of this study which has not been addressed by the authors is the issue of sample purity and how this can significantly affect the calculation of mutant allele frequency. The authors have mentioned that they included samples with >65% tumour cell content and performed micro-dissection on samples with <65%; however, there is still a wide range of TC between 65 - 100% and it is not clear what was the TC in samples with <65% after micro-dissection?

To overcome this issue, couple of approaches can be taken into consideration;

1. Using a computer-based system on scanned H/E slides to calculate the number and percentage of tumour cells in a given sample with high accuracy, and adjusting the mutant AF calculation based on the exact percentage of tumour cells.

OR

2. Performing micro-dissection in all samples in an attempt to purify the samples as much as possible to ideally >90% of TC.

By applying one of the above strategies the calculation of AF would be much more accurate and the corresponding survival data would be more meaningful for practical purposes.

Reply #1: We thank the Reviewer for this suggestion. We fully agree that the estimate of tumor cell percentage is critical for our study. Therefore, as suggested by the Reviewer, two independent expert pathologists have re-calculated the exact percentage of neoplastic cells in

each sample. In order to maintain data accuracy, all cases (n=9) where the exact percentage of tumor cells was not available were excluded. Consequently, the final number of included patients is now 89. To overcome the effects of sample purity on the final results, the obtained variant allele frequency (VAF) for each patient was normalized to the proportion of neoplastic cells in each specimen using the following formula:

$$\text{adjusted VAF (aVAF)} = \frac{\text{VAF}}{\text{TC\%}} \text{ x } 100$$

where VAF represents the percentage of the EGFR variant alleles determined by the pyrosequencing assay and TC% is the estimated percentage of neoplastic cells.

After obtaining the adjusted EGFR-VAF (EGFR-aVAF) for each patient, all statistical analyses were recalculated accordingly. Of note, since the numerical values of EGFR-aVAF were not normally distributed, Mann-Whitney U and Kruskal–Wallis tests were used instead of the originally described Student's t and ANOVA tests, respectively. Importantly, all statistically significant results remained significant after the revision. Our revised data thus further corroborates that EGFR-aVAF is indeed an independent predictor of progression-free survival (PFS) in advanced-stage lung adenocarcinoma (LADC) patients treated with EGFR-TKI.

Changes in the text #1:

- Major revisions were made in the "*Abstract*", "*Methods*", "*Results*", "*Discussion*" and "*Figure legends*" chapters.

- All figures and tables were revised according to the updated results.

Comment #2: Finally, I remain unconvinced, in utilising mutant AF as a robust predictor of response to EGFR inhibitors and I believe all patients with EFGR mutations regardless of the mutant AF measures should be equally treated with EGFR inhibitors, until we have more robust and reproducible data on potential predictive value of this calculation.

The authors should clearly state the latter in the manuscript to avoid any premature interpretation of these data and potential change in practice!!

Reply #2: We thank Reviewer A for raising this point. We agree with the Reviewer that all LADC patients with activating *EGFR* mutations should be equally treated with EGFR inhibitors regardless of the mutant allele frequency, and further analyses are warranted to clarify the predictive role of EGFR-aVAF in these patients (the need for additional studies to validate the results was already mentioned on page 17, lines 384-385). We now clearly state this in the "Discussion" chapter and we also highlighted the need for further studies in the "*Limitations of study*" subchapter.

Changes in the text #2: The following sentences were inserted to the "*Discussion*" chapter:

-page 17, lines 373-376: "Importantly, however, current clinical treatment protocols with regards to EFGR-TKI are still primarily based on the absence or presence of activating EGFR mutations (reference_25). Accordingly, until future validation, the clinicians should choose the most appropriate treatment for their patients regardless of EGFR-aVAF status."

: " All in all, taken into account all the aforementioned potential study limitations, caution is needed when interpreting the results of the present study and further analyses are warranted to clarify the exact predictive role of EGFR-aVAF in EGFR-TKI-treated LADC patients."

-Reviewer B-

We thank Reviewer B for their constructive comments on our manuscript. Below are our answers.

The manuscript by Balazs Gieszer et. al. entitled "EGFR variant allele frequency predicts EGFR-TKI efficacy in lung adenocarcinoma: a multicenter study" aims to evaluate the correlation between EGFR variant allele frequency (VAF) and the sensibility to EGFR TKI. The study was performed on 98 lung adenocarcinoma patients harboring mutated EGFR and the results suggest that high EGFR VAF predicts benefit to EGFR TKI. Although the topic is very interesting, major revisions had to be made before publication

Comment #1: line 93: « Exon 18 mutations are rare and very heterogeneous » I am not sure that these mutations are very heterogeneous since there is still a hotspot on codon 719

Reply #1: We thank Reviewer B for bringing this inaccuracy to our attention. We agree with the Reviewer, that although rare *EGFR*-mutations (such as *EGFR* exon 20 insertions) in NSCLC are generally heterogeneous, exon 18 mutations are indeed relatively homogenous in terms of the hotspot region (1).

Changes in the text #1: The following sentence was revised in the *"Introduction"* chapter:

- page 5, lines 94-97: Exon 18 mutations are rare and relatively homogenous (compared to other rare mutations such as *EGFR* exon 20 insertions) as they represent about 4% of all EGFR mutations (references_9,10).

Comment #2: line 95: authors only cite Gefitinib and Erlotinib, which are first generation EGFR-TKIs. it would be wise to add the 2nd and 3rd generation TKIs.

Reply #2: We thank the Reviewer for picking this up. As suggested, we added the 2nd and 3rd generation EGFR-TKIs to the list of potential therapeutic agents in *EGFR* mutated LADC.

Changes in the text #2: We have modified our text as advised:

- page 5, lines 97-100: Importantly, in LADC, these EGFR-sensitizing mutations confer sensitivity both to first-, second- and third-generation EGFR tyrosine kinase inhibitors (TKIs) such as gefitinib, erlotinib, dacomitinib, afatinib and osimertinib in patients with advanced-stage disease (references_11-13).

Comment #3: line 110 to line 113: this is an interesting study for discussion rather than introduction

Reply #3: We thank the reviewer for this suggestion. The aforementioned sentence was moved to the "Discussion" chapter.

Changes in the text #3: According to this suggestion of the Reviewer, we moved our statement interpreting the finding of Rossi et al. to the "Discussion" chapter:

-page 15, lines 334-336: In addition, based on a recent study on 55 metastatic NSCLC patients, exon 19-mutated patients tend to have better survival outcomes than patients with exon 18 point-mutations as well (reference_15).

Comment #4: line 176: why did the authors decide to exclude cases with less than 65% tumor cells? how was this cut off chosen? This is not what we do on routine practice.

Reply #4: We thank Reviewer B for raising this point. We fully agree that the estimate of tumor cell percentage is critical for our study. Therefore, as also suggested by the Reviewer A, two independent pathologists have re-calculated the exact percentage of neoplastic cells in each sample, and all patients were included to our revised study regardless of the sample's tumor cell percentage. Next, to overcome the effects of sample purity on the final results, the obtained variant allele frequency (VAF) for each patient was normalized to the proportion of neoplastic cells as described in the "*EGFR mutation analysis*" subsection of the "*Methods*" section. Please see Reviewer A answer #1.

Changes in the text #4:

- Major revisions were made in the "*Abstract*", "*Methods*", "*Results*", "*Discussion*" and "*Figure legends*" chapters.

- All figures and tables were revised according to the updated results.

Comment #5: line 183: what is the exact reference of the kit used to search for mutations in the EGFR gene? Therascreen EGFR pyrokit?

Reply #5: We thank the Reviewer for picking this up. Yes, we used the Therascreen EGFR Pyro Kit (ID: 971480, Qiagen) for EGFR mutational analyses.

Changes in the text #5: Clarification regarding the used pyrosequencing kit was added to the "*EGFR mutational analysis*" subchapter.

-page 9, lines 192-194: … using the Therascreen EGFR Pyro Kit (Qiagen) on a PyroMark™ Q24 (Qiagen) pyrosequencing instrument.

Comment #6: The authors never refer to the percentage of tumor cells in the different DNA samples. However, this parameter is very important since it conditions the VAF. This data must be taken into account and the results must be interpreted in function.

Reply #6: We thank Reviewer B for raising this point. As requested, the obtained variant allele frequencies (VAFs) are now normalized to the proportion of neoplastic cells in each specimen. All results, figures and tables were updated accordingly. Please see also our answer to comment #1 of Reviewer A and the revised "*EGFR mutation analysis*" subsection of the "*Methods*" Section.

Changes in the text #6: Clarification regarding the used pyrosequencing kit was added to the "*EGFR mutational analysis*" subchapter.

   - Major revisions were made in the "*Abstract*", "*Methods*", "*Results*", "*Discussion*" and "*Figure legends*" chapters.

   - All figures and tables were revised according to the updated results.

Comment #7: line 223: How to interpret cases with a VAF greater than 70%? Is there an amplification of the EGFR gene? Loss of heterozygosity?

Reply #7: We thank Reviewer B for this comment. Loss of heterozygosity (LOH) and *EGFR* amplification occurs frequently in LADC patients harboring *EGFR* activating mutations, and could serve as an indicator for better response from EGFR-TKI treatment (2-5). Accordingly, as suggested by the reviewer, both of the aforementioned genetic alterations might indeed correlate with higher aVAF values. In the present study, however, we did not investigate the presence of these alterations since they are not part of the routine mutational analyses in Hungary. We now mention this in the last paragraph of *Discussion* describing the limitations of the study.

Changes in the text #7: The following sentences were inserted into the "*study limitations*" subsection of the "*Discussion*" chapter:

   -page 17-18, line 387-394: Loss of heterozygosity and *EGFR* amplification occurs frequently in LADC patients harboring *EGFR* activating mutations and could serve as an indicator for better response from EGFR-TKI treatment (references_44-46). Accordingly, both of the aforementioned genetic alterations might also correlate with higher aVAF values, yet we did not investigate the presence of these alterations since they are not part of the routine mutational analyses in Hungary.

Comment #8: The median EGFR-VAF appears to be very high compared to what is routinely observed in pulmonary adenocarcinomas. This is due to the bias induced by choosing only cases with a percentage of tumor cells greater than 65%. It does not seem to me that this cut off can be applied for other studies or in routine practice.

Reply #8: We agree with Reviewer B that the median and continuous values of the originally included EGFR-VAFs might have been difficult to interpret since initially no normalization was performed. In order to overcome this issue and to bring our findings closer to everyday

practice, all obtained VAFs were normalized to the proportion of neoplastic cells. Thus, we obtained much lower values for EGFR-aVAF that might be hypothesis-generating for further studies. However, we do acknowledge that the newly obtained median value still cannot be used in routine clinical practice. The median value was used solely for the statistical calculations in order to divide the patients into equal subgroups for the Kaplan-Meier approach. Therefore, until further validation, the presented median value can be applied only to our set of patients. For this very reason and to rule out the potential confounding effects of using an arbitrary (yet statistically approved) cut-off, the effects of EGFR-aVAFs on PFS and OS were also measured using the continuous values of the aforementioned parameters (Spearman's correlation). In addition, in the Cox multivariate analysis, we also included the continuous values of EGFR-aVAF rather than using the high vs. low EGFR-aVAF subgroups defined by the median value. Nevertheless, to prevent future misconceptions we now clearly state in the "*Discussion*" chapter that the threshold value to divide patients into low vs. high EGFR-aVAF subgroups was selected based on our dataset, and caution is needed when using it as a cut-off value in future studies.

Changes in the text #8: As recommended by the Reviewer, the following statement was inserted into the Discussion chapter.

-page 16, lines 359-362: It should be noted, however, that the patients were divided into low and high EGFR-aVAF subgroups based on the median value in our dataset, therefore, until further validation, caution is needed when using it as a cut-off value in future studies.

Comment #9: line 352: Authors should add the fact that the study was conducted in selected samples with a high percentage of tumor cells.

Reply #9: We appreciate the Reviewer's suggestion. We agree that our study has some limitations and our findings need to be validated in future studies (as already mentioned in the "Limitations" subchapter [page 17, lines 386-387]). According to the previous suggestions of both Reviewer A and Reviewer B two expert pathologists have re-calculated the exact percentage of neoplastic cells in each sample, and all patients were included to our revised study regardless of the sample's tumor cell percentage.

Changes in the text #9: No changes in the manuscript text.

Comment #10: line 352: another weak point concerns the fact that the whole study was done by evaluating the response to gefitinib and erlotinib whereas today these first generation TKIs are increasingly being replaced by osimertinib

Reply #10: The Reviewer's comment is well taken. We have revised the text of the "Limitations of study" section accordingly.

Changes in the text #10: We have inserted the following sentence to the "Limitations section":

-<mark>page 18, line 392-394</mark>: "Finally, all included patients were treated with first-generation EGFR-TKI erlotinib and gefitinib, yet these inhibitors are being slowly replaced by second- and third-generation EGFR-TKIs in the clinical practice."

Comment #11: Minor points:

1/ Genes should be in italic

2/ The English language in this text would benefit from improvement for clarity and readability

Reply #11: We thank Reviewer B for pointing out this issue. All gene names are now in italic. In addition, a native English speaker has reviewed our manuscript to improve the clarity and readability.

Changes in the text #11:

-All gene names are now in italic.

-Minor grammatical corrections to improve the readability and clarity of the manuscript.

References

1. Beau-Faller M, Prim N, Ruppert AM, et al. Rare EGFR exon 18 and exon 20 mutations in non-small-cell lung cancer on 10 117 patients: a multicentre observational study by the French ERMETIC-IFCT network. Ann Oncol. 2014;25(1):126-131. doi:10.1093/annonc/mdt418

2. Shan L, Wang Z, Guo L, Sun H, Qiu T, Ling Y, Li W, Li L, Liu X, Zheng B, Lu N, Ying J. Concurrence of EGFR amplification and sensitizing mutations indicate a better survival benefit from EGFR-TKI therapy in lung adenocarcinoma patients. Lung Cancer. 2015 Sep;89(3):337-42. doi: 10.1016/j.lungcan.2015.06.008. Epub 2015 Jun 19. PMID: 26141217.

3. Ruiz-Patiño, A., Castro, C.D., Ricaurte, L.M. et al. EGFR Amplification and Sensitizing Mutations Correlate with Survival in Lung Adenocarcinoma Patients Treated with Erlotinib (MutP-CLICaP). Targ Oncol 13, 621–629 (2018). https://doi.org/10.1007/s11523-018-0594-x

4. Hayes DN, McLeod HL. EGFR regulation by microRNA in lung cancer: a rose by any other name ... is an increasingly complicated rose. Ann Oncol. 2008 Jun;19(6):1036-7. doi: 10.1093/annonc/mdn357. PMID: 18508859.

5. Ma ES, Wong CL, Siu D, Chan WK. Amplification, mutation and loss of heterozygosity of the EGFR gene in metastatic lung cancer. Int J Cancer. 2007 Apr 15;120(8):1828-31; author reply 1832-3. doi: 10.1002/ijc.22506. PMID: 17236198.