Reviewer's report

Title: External validation of molecular subtype classifications of colorectal cancer based on microsatellite instability, CIMP, BRAF and KRAS

Version: 0 Date: 11 Apr 2019

Reviewer: Tilman Rau

Reviewer's report:

The study of Alwers et al. is conceptualized as a validation cohort for several proposals of prognostic CRC stratification via molecular characteristics. Particularly, the cohort was split into the combination of subgroups of MSI/MSS, BRAF/KRAS/WT and CIMP. The study is valuable as obviously a detailed monitored follow-up took place and a substantial patient number was reached. However, the strengths of the study are somehow hidden and only appreciated after studying the complete manuscript. Therefore, major revisions should be performed.

1) Abstract and headline need to be sharpened: The main message is the construction and application of an independent external validation cohort to several proposed molecular stratification systems in direct comparison. This is the novel part, but somehow hidden between the lines. It is much more likely that readers see CIMP, MSI and KRAS in colon cancer, which has been answered several times. We deal with competing molecular classification systems in CRC, lack of substantial cohort sizes to address the resulting subgroups etc..

2) The validation cohort as key element: The characterization of the validation cohort is essential. It should be performed according to the REMARK guidelines and integrated into the manuscript as a regular table (not supplement). Although the study was population based, therapies should be outlined, e.g. differences in the application of radiotherapy (relevant for the distinction between CRC and CC), biologicals (particularly cetuximab for the question of KRAS related survival differences as a possible therapeutic effect) and basic information about Chemotherapy yes/no (could be interesting for MSI/MSS differences). Additionally, the prognostics of TNM elements should be shown e.g. in Kaplan-Meier-curves and somehow compared to values of the literature or other epidemiological data sources. A population based validation cohort should reflect and outline the epidemiological and prognostic baseline in CRC today. Additionally, it should be stated why the concept of the population based study called "screening" did not result in younger patients.
3) Molecular findings: As an alternative supplemental table the kind of molecular findings could be precised for all markers, meaning which MMR protein, which KRAS type mutations, differences in number of CIMP loci, frequencies of sporadic MLH1 loss as a somehow logical subgroup MSI/CIMP high etc.. Particularly, the kind of KRAS mutation might explain survival differences as well.

4) Estimation of added values of the different classification systems: A multivariate analysis should show, whether the classification systems would add prognostic value to the TNM-system or not. Additionally, the reader might be interested in performance differences for the three tested systems leading to a recommendation from the point of the authors.

Minor points:

a) How balanced was the inclusion of patients from different sites. Please clarify statistically, whether huge site specific biases might influence outcome data. There should be an exclusion of outliers, if huge differences of surgical/clinical quality could be seen looking at peri-operative mortality or early recidives. This information could be disclosed very generally as information for the reviewer only, but should be checked like in other multi-centric trials.

b) TNM-edition should be mentioned.

c) The selection criteria for the three controlled stratification systems were derived from a review. Please mention them briefly in one sentence.

d) Please explain, why a correlation to the TCGA project was not made.

e) Comment on appropriateness of the TMA approach e.g. comparison IHC versus MSI results.
Are the methods appropriate and well described?  
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?  
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?  
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?  
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal