Author’s response to reviews

Title: Have we achieved adequate recommendations for target volume definitions in anal cancer? A PET-imaging based patterns of failure analysis in the context of established contouring guidelines

Authors:

Hendrik Dapper (hendrik-dapper@gmx.de)
Kilian Schiller (Kilian.Schiller@mri.tum.de)
Stefan Münch (Stefan.Muench@mri.tum.de)
Jan Peeken (Jan.Peeken@mri.tum.de)
Kai Borm (Kai.Borm@mri.tum.de)
Wolfgang Weber (W.Weber@tum.de)
Stephanie Combs (stephanie.combs@tum.de)

Version: 1 Date: 29 Jun 2019

Author’s response to reviews:

A point-by-point response letter

REVIEWER 1:

Reviewer reports:

Julian MM Rogasch, M.D. (Reviewer 1): The authors present interesting and clinically relevant data on the appropriateness or failure of guidelines on CTV definition for RCT in AC. The methodological description, presentation of results and the extent of discussion is generally appropriate. Especially the visual depiction of LNs in Figures 1 and 2 is appreciated.

My concerns and remarks:

General:

I recommend a thorough review of the manuscript regarding specifics of the English grammar (e.g. the correct use of hyphens in words such as "PET-imaging").
In the discussion, I would prefer a less colloquial and/or imprecise language (e.g. page 16, line 44: "In our study we had quite a high number of superomedial […]").

- A specific language correction was performed.

Background:

Page 6, line 54: Please specify what is meant by "patterns of spread". Is it the spread to LN?

- We added “of involved LNs” after “patterns of spread”.

Page 6, line 54: "(CT or MRI)": Please use PET-CT or PET-MRI instead to prevent confusion with the stand-alone imaging modalities CT and MRI.

- We used “PET-CT and PET-MRI” instead of “CT and MRI” alone.

Methods:

Page 8, line 5: Please use the term PET-CT and PET-MRI instead of "CT or MRI based" throughout the manuscript.

- We used “PET-CT and PET-MRI” instead of “CT and MRI” alone throughout the manuscript.

Page 8, lines 30 ff.: Although I understand that the PET imaging protocol is not a central part of the manuscript, a minimal description should be added (injected FDG activity in MBq [with median and range or IQR], interval between injection and start of PET acquisition ["uptaketime"], and the examined field [e.g. base of skull to proximal femora, …]).

- We added details of the PET imaging protocols and reconstruction of the imaging: “Contrast-enhanced FDG-PET imaging was either performed as PET-CT (n = 18; Biograph mCT scanner, Siemens Medical Solutions, Germany) or an integrated whole-body PET-MRI system (n = 5, Siemens Biograph mMR, Siemens Medical Solutions, Germany) after intravenous injection of FDG. In one patient, a PET-CT and a PET-MRI were performed. Median activity of F-18-FDG was 311 MBq (range: 236 – 655 MBq) and the median interval between injection and start of PET acquisition (“uptake time”) accounted for 81 minutes (range: 60 – 108 minutes). The examined field extended from the skull base to the proximal femoral. In one patient with PET-MRI, the detection area included only the abdomen and the pelvis. All patients received oral contrast enhancement. In eleven and four patients additional rectal contrast agent was administered. Twelve of the eighteen patients with PET-CT scan had a diagnostic CT scan of 3
mm slice thickness. In seven cases, low dose CT attenuation correction was needed. The used MRI sequences amounted at least axial/sagittal T2 TSE, axial DWI, axial T1 TSE +/- and sagittal T1. MRI reconstruction was in 3 mm slice thickness. We carried out quantitative evaluation of attenuation-corrected image data by standardized uptake value (SUV) calculation.

PET-CT/MRI reading and interpretation were performed by two experienced nuclear medicine physicians/radiologists. Basically, pelvic LN from 1.0 cm and inguinal LN from 1.5 cm in diameter were considered suspect. However, for the definition of PET-positivity of LN, the combination of different factors such as SUV values, morphology and size of the LN as well as other prognostic factors, such as the tumor stage, were considered.”

Page 8, line 32: 22 patients were included into final analysis, but $18 + 5 = 23$ examinations were performed with PET-CT or PET-MRI. Please check or briefly explain the discrepancy in examination/patient counts.

- We added the information: „In one patient, a PET-CT and a PET-MRI were performed.”

Page 8, line 55: Please correct the typo in "AGIGT"

- We corrected the typo

Page 9, lines 4 ff.: I recommend rewriting this paragraph to make it readily understandable. More specifically, the reader should be able to answer the following questions:

a) Was each patient and each CTV in all patients delineated by each radiation oncologist? Or were different patients/CTV delineated by different readers? Were all three CTV in a specific patient delineated by the same reader?

b) What was the purpose of the template patient? Was it to ensure "training" of the radiation oncologists to ensure uniform understanding of the anatomical regions?

- We mainly rewrote the paragraph: “To obtain an overview of the anatomical distribution of all PET-positive LN of all different patients at the same time, we developed a method to transfer all involved LN on a single CT scan. This was carried out analogously to Schiller et al., who have performed a similar evaluation in prostate cancer [13]. As a first measure, we selected a planning CT scan (3 mm slices thickness) for radiation therapy of a certain AC patient with "standard anatomy" (female, body mass index: 21.7) as a template. Secondly, the three different
CTVs of the current international recommendations were contoured on this CT. The first CTV was defined regarding to the recommendations of RTOG (see figure 1) [8]. The second CTV was delineated analogously to the contouring guidelines of the AGITG and the third to those of the BNG [11,12]. The RTOG and AGITG guidelines for IMRT of AC could be identified via PubMed search using “Contouring guidelines anal cancer”. The BNG is an evidence based consensus for IMRT of AC and currently standard of care within the UK. It is used within the PLATO trial. As the next step, all PET-positive LN of the twenty-two patients were delineated on the one chosen CT scan (template) by an experienced radiation oncologist. To transfer the LN to the template as accurately as possible, the anatomical conditions of each positive LN in the original PET imaging of all twenty-two patients were considered (relations to e.g. vessels or musculoskeletal structures). LN locations were defined as inguinal, external and internal iliac (including obturator nodes), pre-sacral, para-rectal, common iliac and para-aortic, and were recorded in a table (table 1). The LN were contoured by standard starting from the centre of the LN consistently on three axial CT slices (longitudinal extension: 9 mm) by using a brush with 9 mm diameter to represent each LN at 9 x 9 mm. Afterwards, the radiation oncologist evaluated whether these LN were covered by the three CTVs of the different contouring guidelines. This was done individually for each of the three CTVs. The definition of “miss” arose from the fact that the majority (>50%) of the volume of the LN was not covered by the CTV. Using a color code, the LN metastases were divided indicating whether their location was in- (green) or out-field (orange) of the standard CTV. The process of LN transfer to the template and the decision as to whether a LN was predominantly included within a particular CTV, was reviewed by at least one other experienced radiation oncologist.”

Results:

Page 10, line 16: Please consider using table 1 instead of 2 as the reference for the 154 LNs.
- We changed the reference to table 1.

Page 11, line 57: Please explain that "(14/7/5)" refers to the three different guidelines, e.g. "(RTOG: 14; AGITG: 7; BNG: 5)".
- We changed the sentence as you recommended to “(RTOG: 14; AGITG: 7; BNG: 5)”.

Page 12, lines 1 ff.: In the context of logistic regression (or other methods that are not "true" correlation methods), I recommend avoiding the term "correlation" not to confuse two different statistical methods. Please consider terms like "association" or "relationship". Generally, I recommend reconsidering the choice of statistical method. As I understand, the authors aim at
evaluating merely the association between a single clinical variable (e.g. T stage) and the LNs being inside or outside of the CTV. The combination of several factors in one model (e.g. T stage and grading) appears dispensable. For this purpose, a chi square test applied to each clinical variable separately should be the appropriate (and simpler) test. A logistic regression would only be necessary if the aim was to test and model the predictive significance of several variables (T stage, grading, ...) in a combined model. However, this would imply a more differentiated description of the model (choice of independent variables) in the methods section, a more complex description of the results (overall accuracy of the model, details on the regression coefficients of the variables) and a different terminology in the results description ("prediction" instead of "association"/"relationship"). Furthermore, considering the limited sample size and limited selection of independent variables (only T stage and grading!), the reliability and power of a method such as a binary logistic regression seems generally questionable.

- As you mentioned, due to the fact that these statistical studies on the actual topic of the paper are rather unimportant, we have modified the statistics according to your recommendations.

- Changed Method: “Statistical analysis was conducted using ‘IBM SPSS statistics’ software, version 23.0 (IBM, Armonk, USA). A Chi-Square test was applied to analyse differences regarding T-stage and the distribution of LN outside or inside the CTV of RTOG.”

- Changed Results: “Dependent on the T-stage, LN showed a significantly different distribution of being outside or inside of the CTV of RTOG”.

Discussion:

Page 17, line 1: Please correct the typo in "leed"

- We corrected the typo “leed” to “lead”

Page 18, line 26: For a reader who is not experienced in radiation oncology, it may be difficult to understand the concept of transferring the "positive LNs of 22 patients on one patients' planning CT scan", especially if this appears to result in distortion of anatomical marks. In analogy to my initial remark, I recommend describing the benefit or aim of the one patient's planning CT in the framework of the current study (especially considering that some readers may be unexperienced in radiation oncology).

- We have significantly revised and explained the paragraph within the methodology and especially explained the concept of transferring the "positive LNs of 22 patients on one patients' planning CT scan. Please see the chapter “Methods”.
"Limitations" paragraph: I agree with the authors that the detected LNs and the misses have to be interpreted considering the overall tumor spread (and thus M0 / M1 situation and risk for further LN metastases). However, I recommend specifying or clarifying the discussion on two issues:

a) What is the guidelines' general application of "elective" CTV? What is the clinical framework (e.g. cT and cN stage) that implies the appropriateness of an elective CTV instead of an individually confined / extended CTV?

b) To which degree do all 22 patients in the current study and, more importantly, the patients with LN misses, comply with the guidelines' framework of an "elective" CTV use? This could include a brief statement on the in-/exclusion criteria (patients with common iliac or para-aortic LN were not excluded).

- We extend the discussion/Limitations by the following paragraph: “Special care must be taken when interpreting the inguinal misses in patients with extensive loco-regional situations which are defined as metastatic disease (M1, LYM). We have included these patients as they were treated in curative intention with a standard protocol of CRT. These cases are not representative and therefore basically not useful to derive a meaningful elective CTV definition for all patients. In principle, guidelines serve as orientation for a reasonable standardized target volume in order to cover potential micrometastases and to save regions with very low risk of tumor invasion to reduce toxicity. Of course, these prescriptions are abandoned in real clinical scenarios when macrometastases appear in the imaging (e.g. para-aortic LNs would be included). However, these loco-regional advanced cases provide fundamental reference for possible anatomical patterns of inguinal involvement as it can be assumed that some of these PET positive LNs were already affected but not visible or morphologically suspicious at an earlier point in time with clinically lower stage.”

Page 18, line 28: Please correct the typo in "seize"

- We changed the typo in “size”.

REVIEWER 2:

Shiro Watanabe (Reviewer 2): Comments for manuscript, BCAN-D-19-00907
The authors examined the difference of 3 contouring guidelines for the CTV in AC with FDG PET imaging as reference. LN "misses" appear generally cranially or caudally to the recommended CTVs. The established guidelines differ significantly in the inguinal region, and authors suggested for CTV definition of the inguinal region.

This is an important retrospective study evaluating pattern of LN spread and comparing the contouring guidelines in detail.

Although a small sample size, this does lay the groundwork for future studies. Overall, I think it reasonably well written, however, I have several specific comments and suggestions.

Specific comments

Background

1. P6 L10, P6 L29:

I could not confirm the cited content in [3] and [8]. Please correct.

- Reference 3: Thank you! We chose the wrong citation. Now we changed the references.
- Reference 8: This content is right and can be found at pubmed (PMC full text).
  https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2709288/

Method

2. P7 L31:

Details regarding FDG PET and CT or MRI acquisitions are lacking. FDG PET/CT and FDG PET/MRI protocol and reconstruction needs to be addressed.

- We added details of the PET imaging protocols and reconstruction of the imaging.

3. P7 L40:

The definition of PET-positivity was not explained well. Although FDG PET imaging is very sensitive for LN metastasis, false positive/negative is not uncommon. The reproducibility remains unclear. The criteria for PET-positivity need further specification and explanation, so
that the reader can follow the justification for the grouping. If needed, please comment in the limitation section.

- As there is currently no evidence for clear SUV\text{max} cut off values, we described our rules as the following:

  - We added “Basically, pelvic LN from 1 cm and inguinal LN from 1.5 cm in diameter were considered suspect.” to the method section.

  - We added the following sentences to the limitations: “There were no strict and thus reproducible criteria by which LNswere finally classified as involved. A certain degree of uncertainty (false positive/negative) is, however, inevitable since a final assessment must always take various factors into consideration.”

4. P9 L14:

Please explain the contouring method if LN was smaller than 9mm.

- The LNs were standardized contoured starting from the centre of the LN consistently on three axial CT slices (longitudinal extension: 9 mm) by using a brush with 9 mm diameter so that each LN was represented at 9 x 9 mm

Results

5. P10 L6:

Characteristics of 22 patients should be given in Table.

- Since the maximum number of tables has already been reached, we have integrated the T-stage as a tumor characteristic into Table 1.

6. P10 L16:

table 2 → Table 1?

- We changed table 2 with table 1.
Discussion

7. P13 L6:

153 → 154?

- Thank you, we changed the number of LNs to 154.

8. P18 L15:

Please add the number or percentage instead of just "reasonable number". I have the question whether LN "misses" in inguinal regions really contribute to low local control rate in CRT.

- We added the number of patients and percentage.

Tables

9. P26 L1:

This sentence is copy of figure 3. Please correct.

- Thank you. We corrected that mistake.

Figures

10. Figure3:

Please add SUV color scale bar.

- We added the information of SUVmax and SUVmean values of the three different LN in the description of figure 3.