

Tumor control and immune activation through palliative irradiation and ATR inhibition, PATRIOT Part C: a phase Ib trial

Corresponding Author: Dr Magnus Dillon

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The manuscript by Dillon et al. reports the results of an early phase clinical trial evaluating the safety, efficacy and translational correlates of an ATR inhibitor with low-dose palliative radiotherapy. They document a few exceptional responses with durable long-term control to this treatment and highlights its possibility for promising future development, which is evident with subsequent ongoing studies. These are important results that will help shape the development of this therapeutic combination in ongoing and future clinical trials.

That being said, a number of key results are poorly described and hard to interpret and the results needs significant improvement to help with clarity. Specific comments follow below:

- The review of DLTs and Gr3 toxicities is quite confusing and hard to follow. For instance, on Pg4, line 1-3. It is unclear how many patients are evaluable for safety results – it sounds like 5 patients who had toxicities were not included, although Figure 1D says these were included. Please clarify this. Further, its not quite clear exactly how many people had Grade 3 events. For instance in the H&N 8/8 3 wk 30Gy cohort, it seems like there were 9 Grade 3 toxicities – but its unclear how many unique patients this occurred in. Presumably this arm was deemed to be too toxic to proceed? The text describing the toxicities is quite convoluted – just documenting individual events without providing any type of summary or structure. Some type of summary of what was considered a DLT should be provided in the main text and shouldn't require review of supplementary material to figure out. The rate of Grade 3 toxicities with low doses here seems a little high?
- Figure 1E itself is quite hard to interpret as it shows a number of OARs that have no side effect. It also shows multiple metrics for some organs (lung) which are likely highly correlated. This figure would be much better if it was simplified to focus on OARs that have side effects as a main figure; with the full figure reserved as a supplement. The text describing the figure also doesn't really provide any type of take home message – although presumably the authors are trying to suggest there may be a volume effect leading to Gr3 toxicities in some organs?
- Figure 2B – why does F2 time point look like higher levels of pChk1 than time point F1 – is this from RT inducing some signaling?
- Responses in un-irradiated lesions are hard to assess – it seems like Sup Fig 7 & 8 may have included them as target lesions? Since we have responses for irradiated lesions; would it be possible to see responses in unirradiated lesions separately? The authors say “there were no abscopal responses” This suggests unirradiated lesions had no response – but this is not clear from any of the data provided.
- The hypopharynx patient who had long term disease control – was his tumor p16+ or HPV-related? More detail should also be provided about the maxillary adenoca patient who had 3 year disease control apparently.
- In the 5 of 22 patients that developed progression of disease, what was time to POD.

- Why is median follow up only 158 days for a trial that completed in 2020? Some information on overall survival should be provided (just to get a sense of how sick the cohort was).

- Its not clear what can be taken away on paired RNA-seq from 2 cases 4C-4F. This analysis to me doesn't seem to add much. Pg 6 line 13-16 summary statement seems a bit too strong for the data present .

- The discussion would benefit from putting immunologic results in context of recent work from Chris Bakinest's group on toxicity of ATR to T-cells, and work from AZ on the myelodepletive effects of AZ in the TME. Further recent results on pre-clinical work showing genotype specific radiosensitization in ATRi merit at least some discussion in context of trial results here.

Minor

- Table 1 should include how many patients had metastatic vs. loco-regional only disease
- Supp Table 1 only 26 patinets are included in RT techniques of 27 patinets presumably treated
- In evidence of target modulation – figure 2D is not referenced; there is also small error on line 49, pg 4, with 1E presumably meant to be 2D
- Pg 5, line 5-9 – please make it clear this is response in irradiated lesions in text.
- Figure 4A – please indicate type of sequencing in oncoprint used for calls

Reviewer #2

(Remarks to the Author)

1. A 3+3 design was used for part C. If I understand the design correctly, patients were assigned to either C1 or C2 depending on their irradiation site. In both C1 and C2, the first three patients were enrolled in a dose cohort (DL1). If no DLTs occurred, the study progressed to the next higher dose cohort (DL2). I had a hard time to understand how patients were enrolled to the trial based on the design. I was wondering if authors could provide a CONSORT type of diagram to show a flowchart that visually depicts the flow of participants through each patient and dose cohort of the trial (It would be better to reorganize Fig 1D).

2. 18.5% (5/27) patients were excluded from the study because 4 patients did not complete 1 scheduled treatment due to toxicity and 1 patient had disease progression before response assessment. However, these events occurred after the start of treatment and were related to treatment. They were like informative censoring to observe the patient's response and therefore excluding these patients from the analysis may introduce a bias in the estimation of the effect of treatment.

3. In the main body text, CR, PR, PD, etc should be reported with 95% CI

4. Due to the size of the sample size, IQR would no be a good summary statistic (for example Table 1).

5. In Table 1, "unknown" was excluded to calculate the proportion in both blow and above, but included in the total. There is inconsistent in handling "unknown". Also no percentages were displayed for any previous radiotherapy and primary diagnosis by diaphragm.

6. I was wondering if the analyses reported in Figure 5 were performed correctly. The comment mentioned Fold-change from baseline in percent CD69-positive cells was performed by Wilcoxon test (also D, and H). Is Wilcoxon test is Wilcoxon signed rank test or Wilcoxon rank sum test? The null hypothesis of Wilcoxon signed rank test is the median of the population of differences between the paired data is zero, and that of Wilcoxon rank sum test is two independent samples are drawn from populations with the same distribution. What does "A hypothetical median of 1" mean? No multiplicity was considered in the analyses. If multiplicity is considered, almost results were not significant. For G, the data was repeatedly collected. I just wonder if the repeated measure analysis was considered for the analysis? What does 2-way ANOVA mean?

Reviewer #3

(Remarks to the Author)

Dillon et al. present their findings on a clinical trial investigating the safety and tolerability of concurrent ATR inhibitor (here ceralasertib) with palliative RT therapy. This clinical trial separately investigated ceralasertib as monotherapy, and the authors report here Part C of the study, which was further divided into 2 study arms according to whether the site of irradiation was above or below the diaphragm. There were 3 dose levels for ceralasertib and 2 dose levels for RT.

This is a very interesting study, investigating a hypothesized interaction between DDR, which has drugs employed largely as systemic therapy currently in combination with a locoregional therapy – palliative RT.

Further discussion on several points would improve this manuscript.

Page 4 Line 11 and 12 refer to evidence of radiosensitization in the AE profile of normal tissues as justification for the highest DL of ceralasertib. Given this important decision, perhaps Figure 1 should be referred to and briefly expanded upon.

Page 5 Lines 5 refers to tumor growth kinetics before entering the study. Where these all in the context of active therapy? In the corresponding figure, much of these tumors appear to have been minimally increasing, were these the that were progressing? Or was it the out-of-field lesions? A little more background on the analysis underlying this claim would be helpful.

Page 5 Line 18, what separate malignancy? Is there a concern for a distal safety signal here?

TP53 is mentioned as a gene of interest as per Page 6 Line 12. The supp GSEA figure does not appear to have TP53 pathway upregulated in it though. KRAS is also discordant. Are these two truly analogous?

Page 7 line 4 mentions the durable responses are greater than would be expected with palliative radiotherapy alone. But then in line 18-20 there is mention that there is a lack of proper reference baseline studies. These appear to be contradictory claims. More discussion would be helpful or otherwise avoiding a direct comparison.

The changes in localized immune responses with temporal data is very interesting and as the authors point out worth additional studies. The hypothesis as pointed out in page 8 lines 4-6 is interesting, but perhaps a little strongly stated at current.

Overall, the discussion and analysis appear to collapse the 2 study arms above/below the diaphragm into one study group to report results. It would be helpful to clarify throughout the manuscript or otherwise comment on why the study arms are not discussed or analyzed separately.

Reviewer #4

(Remarks to the Author)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have addressed all of my prior concerns thoughtfully and comprehensively. I have no remaining reservations and recommend the manuscript for publication. Congratulations to the authors on their excellent work.

Reviewer #3

(Remarks to the Author)

The authors have reviewed their manuscript and provided a great more clinical detail on individual patient cases as suggested by the other reviewers. This is helpful and has improved the manuscript. The additions to Figure 1 are also very welcome.

We are satisfied with the author's responses to our comments and the corresponding revisions within the manuscript.

Reviewer #4

(Remarks to the Author)

The authors have revised their manuscript and provided a great more clinical detail on individual patient cases as suggested by the other reviewers. This is helpful and has improved the manuscript. The additions to Figure 1 are also very welcome.

We are satisfied with the author's responses to our comments and the corresponding revisions within the manuscript.

Reviewer #5

(Remarks to the Author)

I am not the original reviewer 2, but I was asked to review the authors' responses to reviewer 2's comments. I think the authors did a great job of responding to all of reviewer 2's comments, except for their response to the 3rd comment. I think the authors may have misunderstood this comment. Reviewer 2 did not ask for mean/std etc., but for 95% confidence intervals for these rates. This is a reasonable comment and not difficult to address. For example, the authors could use binomial (Clopper-Pearson) methods based on the number of evaluable participants. These changes will improve the interpretability and rigor of the response data.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Reviewer #1:

The manuscript by Dillon et al. reports the results of an early phase clinical trial evaluating the safety, efficacy and translational correlates of an ATR inhibitor with low-dose palliative radiotherapy. They document a few exceptional responses with durable long-term control to this treatment and highlights its possibility for promising future development, which is evident with subsequent ongoing studies. These are important results that will help shape the development of this therapeutic combination in ongoing and future clinical trials.

That being said, a number of key results are poorly described and hard to interpret and the results need significant improvement to help with clarity. Specific comments follow below:

- The review of DLTs and Gr3 toxicities is quite confusing and hard to follow. For instance, on Pg4, line 1-3. It is unclear how many patients are evaluable for safety results – it sounds like 5 patients who had toxicities were not included, although Figure 1D says these were included. Please clarify this.

We thank the reviewer for these observations. From the very outset, we appreciated the central importance of safety and toxicity reporting when combining radiotherapy with a DNA damage response inhibitor. Therefore, toxicity data is included for all participants who were dosed (n=27). Where patients did not complete the requirements for treatment administration to be eligible for DLT assessment, they were replaced in order to complete cohorts before any dose-escalation decisions could be taken by the safety review committee. This has been clarified in page 3 line 35-41, first paragraph of results section:

“Twenty-seven patients were dosed between January 2016 and December 2020 and included in the safety analysis (6 in the below-diaphragm cohort and 21 in the above-diaphragm cohort). Of these, 22 were considered evaluable for DLT (completed at least 80% of planned treatment) and 23 for response (completed at least one response assessment) (Fig. 1D). Four did not complete planned treatment and were replaced, 2 due to grade 3 thrombocytopenia, 2 due to patient choice with low grade toxicities. A further patient progressed clinically before the first response assessment.”

Further, it's not quite clear exactly how many people had Grade 3 events. For instance in the H&N 8/8 3 wk 30Gy cohort, it seems like there were 9 Grade 3 toxicities – but it's unclear how many unique patients this occurred in. Presumably this arm was deemed to be too toxic to proceed? The text describing the toxicities is quite convoluted – just documenting individual events without providing any type of summary or structure. Some type of summary of what was considered a DLT should be provided in the main text and shouldn't require review of supplementary material to figure out. The rate of Grade 3 toxicities with low doses here seems a little high?

We apologise to the reviewer for presenting the data in a convoluted way and we hope that these data are now clearer. We have added a row for each dose cohort specifying the number of unique patients who experienced the indicated grades of toxicity. In fact, the majority of the G3 toxicities in the highest dose level are in the

same patient who experienced a DLT. In the 2 other participants, the toxicities were (i) G3 leukopenia (ii) G3 anaemia and leukopenia. We have shaded the columns to indicate the high-grade toxicities and the DLTs which occurred in the same patient, added a statement about DLT definition (page 3, line 45-46), and re-worked the text as follows to make it clearer (page 3, line 35 to page 4, line 7):

“Twenty-seven patients were dosed between January 2016 and December 2020 and included in the safety analysis (6 in the below-diaphragm cohort and 21 in the above-diaphragm cohort). Of these, 22 were considered evaluable for DLT (completed at least 80% of planned treatment) and 23 for response (completed at least one response assessment) (Fig. 1D). Four did not complete planned treatment and were replaced, 2 due to grade 3 thrombocytopenia, one due to patient choice with low grade toxicities and one for unknown reasons. A further patient progressed clinically before a the first response assessment. Administered radiation dose was 20 Gy in 10 fractions in 13 participants and 30 Gy in 15 fractions in 9 patients. Baseline characteristics are given in table 1. Radiation details are in Supp. Table 1.

DLT was defined as G4 hematological, G3 non-hematological (see supplementary data for full details) toxicity, G3 or G4 radiation-induced skin or mucosal toxicity or the inability to tolerate one week of combination therapy due to toxicity. There was one DLT at the highest administered dose of 80 mg BD ceralasertib and 30 Gy in 15 fractions RT with all RT fractions sensitized. The DLT was grade 3 mucositis and dermatitis, causing dysphagia, hypokalemia and hypophosphatemia requiring hospital admission in a patient treated for oropharyngeal squamous cell carcinoma.

Adverse events felt possibly, probably or definitely related to ceralasertib and/or radiotherapy are listed in table 2. The profile of toxicities was mainly hematological. There were mild radiation reactions in the majority of participants, with G1-2 dermatitis, mucositis, esophagitis observed, but these were not felt to be significantly greater than would be observed with radiation alone. At 30 Gy with 3 weeks' ceralasertib, 3/8 participants had G3 toxicity, however only 1 participant had G3 radiation sensitization (DLT patient above), with the remainder experiencing hematological toxicity. Non-DLT SAEs related to the therapy were G3 radiation-induced mucositis and dermatitis (the DLT at 80 mg/30 Gy), which was associated with G3 dysphagia and hypokalemia, and one episode of febrile neutropenia, with G1 neutropenia and thrombocytopenia (80 mg/30 Gy).”

ceralasertib dose/ radiation dose	20mg/20Gy		40mg/20Gy	40mg/20Gy	80mg/20Gy	80mg/20Gy	80mg 2 wk 30Gy	80mg 3 wk 30Gy	total (number of unique patients)				
									n=27				
evaluable/total	H&N n=3/4		A&P n=3/3	H&N n=3/3	H&N n=2/3	A&P n=1/3	H&N n=3/3	H&N n=8/8					
CTCAE Term	G<3 No.	G≥3 No.	G<3 No.	G<3 No.	G<3 No.	G<3 No.	G<3 No.	G<3 No.	G≥3 No.	G1-4 No.	%	G3-4 No.	%
Dermatitis radiation				2	1		2	7	1	13	48.1	1	3.7
Fatigue					1		1	4		6	22.2	0	0.0
Anemia		1	1	1			1	2	1	7	25.9	2	7.4
Mucositis oral				1	1		1	2	1	6	22.2	1	3.7
Nausea	1					1	1	2		5	18.5	0	0.0
Thrombocytopenia		1					1	2		4	14.8	1	3.7
Pain					1			1	1	3	11.1	1	3.7
Vomiting			2			1				3	11.1	0	0.0
Anorexia							1	2		3	11.1	0	0.0
Diarrhea		1	2							3	11.1	1	3.7
Esophagitis	1			1						2	7.4	0	0.0
Skin infection				1				1		2	7.4	0	0.0
Dizziness								2		2	7.4	0	0.0
Hypotension								2		2	7.4	0	0.0
Odynophagia								2		2	7.4	0	0.0
Dysphagia								1	1	2	7.4	1	3.7
Leukopenia									2	2	7.4	2	7.4
Hypokalemia									1	1	3.7	1	3.7
Hypophosphatemia									1	1	3.7	1	3.7
number of unique patients	2	3	3	2	2	1	3	7	3	22	80	6	20

Table 2: TRAE related to radiotherapy, by dose cohort

Number in each cohort refers to the total number of participants who experienced at least 1 AE/the total number of patients assessable for toxicity; AE numbers in rows are for AEs which were experienced by >1 participant or were G ≥ 3. Total columns express total number out of 27 participants evaluable for toxicity. Blue cells indicate the DLT, all in the same participant.

- Figure 1E itself is quite hard to interpret as it shows a number of OARs that have no side effect. It also shows multiple metrics for some organs (lung) which are likely highly correlated. This figure would be much better if it was simplified to focus on OARs that have side effects as a main figure; with the full figure reserved as a supplement. The text describing the figure also doesn't really provide any type of take home message – although presumably the authors are trying to suggest there may be a volume effect leading to Gr3 toxicities in some organs?

We appreciate this suggestion. The multiple metrics were included to aid future studies of this combination to determine which standard dose-volume constraints are safe in our experience, in order to design sensible dose-volume constraints in future. We have tried to make clearer the fact that the take-home message regarding toxicity in the lung (G3 event) is that a significant volume of lung was irradiated. We have emphasised this point in the following text (page 4, line 21). We have simplified the graph, as suggested, and present a version that is more readily accessible in the main paper, while retaining the entire data set by moving the full figure to supplementary data.

“One patient being treated for a lung tumor (with 30 Gy and 80 mg BD, 10 fractions sensitized) developed CTCAE G2 pneumonitis at 3 months after the start of

radiation, which responded to steroids. A significant volume of lung was irradiated, with lung V_{20Gy} was 33.5% and mean lung dose 11.3 Gy.”

- Figure 2B – why does F2 time point look like higher levels of pChk1 than time point F1 – is this from RT inducing some signaling?

Yes, the reviewer is correct. RT induces Chk1 signalling as shown in our preclinical studies (DOI: 10.1158/1535-7163.MCT-16-0239). This makes the effect of ATRi+RT difficult to interpret in this context.

- Responses in un-irradiated lesions are hard to assess – it seems like Sup Fig 7 & 8 may have included them as target lesions? Since we have responses for irradiated lesions; would it be possible to see responses in unirradiated lesions separately? The authors say “there were no abscopal responses” This suggests unirradiated lesions had no response – but this is not clear from any of the data provided.

Un-irradiated lesions were included as target lesions if they were present. Hence, we present response data as ‘overall’ including all measured lesions, and ‘irradiated only’ because, in many cases, out of field lesions continued to progress despite response in the radiation field. Given the mechanism of action of the drug, we agree with the reviewer that presenting the unirradiated lesions is important, and more valuable than the sum of all lesions. We have added these figures to supplementary Fig. 8 and 9. During this process we identified some misclassification of irradiated and unirradiated lesions, which is why the CR and PR rates for the irradiated lesions have been corrected through the manuscript.

- The hypopharynx patient who had long term disease control – was his tumor p16+ or HPV-related? More detail should also be provided about the maxillary adenoca patient who had 3 year disease control apparently.

Regrettably, because this patient’s tumour was in the hypopharynx, it was not stained for p16 by immunohistochemistry or subjected to in situ hybridisation analysis for HPV. Therefore, its status is recorded as HPV unknown, although we have noted his heavy smoking history and subsequent death, some years later, from a second primary lung cancer. This important information has been added in the manuscript (page 5, line 12). We did run the Oncoanalyser pipeline on the DNA sequencing data from the tumour, which includes this viral detection tool:

https://github.com/PapenfussLab/gridss/blob/master/VIRUSBreakend_Readme.md ;

It did not show any sign of HPV, but this tool would only have been useful if the result was positive, so we have not included this information in the revised manuscript.

- In the 5 of 22 patients that developed progression of disease, what was time to POD.

Given the fact that these patients were treated with palliative intent, the protocol did not include a rigorous schedule of follow-up imaging investigations that would allow us to declare accurate progression-free or event-free survivals. As such, the progression dates are not available for all participants. Instead, we have presented the overall survival, which is a more certain endpoint and available in all cases. This information has been added to the manuscript (page 5, line 28) and is presented in supplementary Figure 11.

- Why is median follow up only 158 days for a trial that completed in 2020? Some information on overall survival should be provided (just to get a sense of how sick the cohort was).

The median overall survival is, of course, influenced by the fact that many patients received this treatment as an end-stage palliative care intervention. We have added information on overall survival in Supp. Fig. 11, referenced in page 5, line 28.

- Its not clear what can be taken away on paired RNA-seq from 2 cases 4C-4F. This analysis to me doesn't seem to add much. Pg 6 line 13-16 summary statement seems a bit too strong for the data present .

We accept the fact that we are only able to present very limited numbers of paired biopsy samples. However, given the rarity of being able to sample patients receiving an experimental drug and palliative RT, we believe that these data represent an important addition to the literature. In line with the referee's request, have toned down some of our interpretation of the data, but we have retained the figure and the discussion of its relevance: page 6, line 12-13:

"Taken together, these data support further studies of the microenvironmental response to ATRi-RT, in particular early stress responses, innate immune activation and adaptive immune cell depletion, before a subsequent repopulation with adaptive immune cells."

- The discussion would benefit from putting immunologic results in context of recent work from Chris Bakinest's group on toxicity of ATR to T-cells, and work from AZ on the myelodepletive effects of AZ in the TME. Further recent results on pre-clinical work showing genotype specific radiosensitization in ATRi merit at least some discussion in context of trial results here.

We agree with this comment and have added reference to the work that is mentioned in the discussion section:

- Page 7, line 18-20: genotype-specific radiosensitisation (Ng et al.)
- Myeloid effects, page 7, line 55 (Hardaker et al.)
- ATRi toxicity to T cells, page 8, line 4-6 (Vendetti et al.)

Minor

- Table 1 should include how many patients had metastatic vs. loco-regional only disease

We have added this important information.

- Supp Table 1 only 26 patinets are included in RT techniques of 27 patinets presumably treated

We thank the reviewer for pointing out this oversight and have added the missing patient whose RT technique was not recorded.

- In evidence of target modulation – figure 2D is not referenced; there is also small error on line 49, pg 4, with 1E presumably meant to be 2D

Thank you for spotting these, they have been corrected.

- Pg 5, line 5-9 – please make it clear this is response in irradiated lesions in text.

This has been re-worded for emphasis. Page 5, line 1:

“In irradiated lesions, of 23 evaluable participants (who had at least 1 response assessment), 2 (9%) had RECIST complete response (CR), 6 (26%) partial response (PR), 13 (57%) stable disease (SD) and 2 (9%) progressive disease (PD),”

- Figure 4A – please indicate type of sequencing in oncoprint used for calls

We have updated this figure with the requested information. The top of the oncoprint ‘sample_type’ tells us which sample type has been used. Plasma sequencing, using a targeted exome panel, or tumour sequencing using either whole exome sequencing or a targeted exome panel was used. We have changed the oncoprint to add in this information and updated the figure legend of Fig. 4A to clarify this.

Reviewer #2:

1. A 3+3 design was used for part C. If I understand the design correctly, patients were assigned to either C1 or C2 depending on their irradiation site. In both C1 and C2, the first three patients were enrolled in a dose cohort (DL1). If no DLTs occurred, the study progressed to the next higher dose cohort (DL2). I had a hard time to understand how patients were enrolled to the trial based on the design. I was wondering if authors could provide a CONSORT type of diagram to show a flowchart that visually depicts the flow of participants through each patient and dose cohort of the trial (It would be better to reorganize Fig 1D).

We are grateful to the reviewer for pointing out these difficulties in interpreting our original figure. We have attempted to clarify fig 1D, indicating patient flow in the study. The reviewer is correct that a 3+3 design was used. More detail has been added on the 3+3 design in the methods section. The lack of numbers in the below diaphragm cohorts reflects poor recruitment rather than study design.

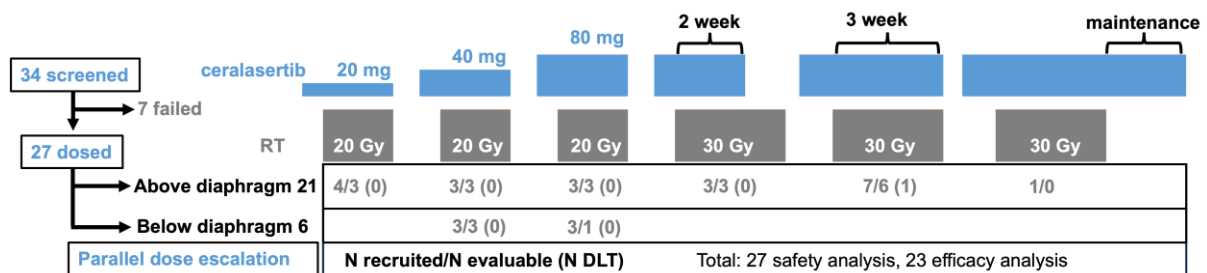
Page 2, line 27-32:

“This was a multi-center, open-label, phase I dose-escalation study, using a 3+3 design (Fig. 1A), where 3 participants were recruited per dose cohort, with dose escalation if there were no dose-limiting toxicities (DLT), and expansion to 6 participants if there was 1 DLT, with the maximum administered dose defined as the dose level at which 2 participants experienced a DLT, and maximum tolerated dose the dose level below, where no more than 1/6 participants had DLT.”

Page 2, line 35-36:

“The data from both arms have been presented together, due to no clear site-specific DLT and poor recruitment below the diaphragm.”

Fig. 1D



2. 18.5% (5/27) patients were excluded from the study because 4 patients did not complete 1 scheduled treatment due to toxicity and 1 patient had disease progression before response assessment. However, these events occurred after the start of treatment and were related to treatment. They were like informative censoring to observe the patient's response and therefore excluding these patients from the analysis may introduce a bias in the estimation of the effect of treatment.

We have clarified these points in the text (page 3, line 35-41). We can reassure the reviewer that **ALL** patients who were dosed were included in the toxicity analysis, which is the primary endpoint of the study. We have changed the responses quoted to reflect all participants who were dosed and received at least 1 response assessment.

Updated page 3, line 35-40

"Twenty-seven patients were dosed between January 2016 and December 2020 and included in the safety analysis (6 in the below-diaphragm cohort and 21 in the above-diaphragm cohort). Of these, 22 were considered evaluable for DLT (completed at least 80% of planned treatment) and 23 for response (completed at least one response assessment) (Fig. 1D). Four did not complete planned treatment and were replaced, 2 due to grade 3 thrombocytopenia, one due to patient choice with low grade toxicities and one for unknown reasons."

3. In the main body text, CR, PR, PD, etc should be reported with 95% CI

We have discussed this and, with respect, do not feel that this is an appropriate form in which to present the response data as we are not using averages to describe the response percentages.

4. Due to the size of the sample size, IQR would no be a good summary statistic (for example Table 1).

We accept this comment and have removed IQR from table 1

5. In Table 1, "unknown" was excluded to calculate the proportion in both blow and above, but included in the total. There is inconsistent in handling "unknown". Also no percentages were displayed for any previous radiotherapy and primary diagnosis by diaphragm.

Thank you for this observation, these errors have been corrected.

6. I was wondering if the analyses reported in Figure 5 were performed correctly. The comment mentioned Fold-change from baseline in percent CD69-positive cells was performed by Wilcoxon test (also D, and H). Is Wilcoxon test is Wilcoxon signed rank test or Wilcoxon rank sum test? The null hypothesis of Wilcoxon signed rank test is the median of the population of

differences between the paired data is zero, and that of Wilcoxon rank sum test is two independent samples are drawn from populations with the same distribution. What does “A hypothetical median of 1” mean? No multiplicity was considered in the analyses. If multiplicity is considered, almost results were not significant. For G, the data was repeatedly collected. I just wonder if the repeated measure analysis was considered for the analysis? What does 2-way ANOVA mean?

The data presented in Fig 5C-E expressed the proportion of cells expressing the indicated marker normalised to the baseline sample for each patient. Hence, the baseline measure is defined as 1. We tested each result for normal distribution using the Shapiro-Wilk test. For normally-distributed data, we then used a 1-sample t-test and, for data that was not Gaussian, we used a Wilcoxon signed-rank test (this has been clarified in the figure legend). Both of these tests compare the data against a hypothetical median, which in this case is 1. We did this because that is the normalised figure for the baseline value and we wanted to see if the on-treatment results deviated significantly from baseline. These data vary significantly by patient at baseline, so we felt that the normalised data is the most appropriate form in which to present the information. We have changed 5C to reflect statistically significant findings with multiplicity testing, and changed the associated text to reflect the data presented: page 6, lines 30-44

For G, the reviewer is correct that the data represents repeated collections from the same subjects. 2-way ANOVA was used to compare the 2 columns (response vs. non-response). Repeated measure analysis was not used. We have changed the figure and associated text to remove the significance values (page 6, line 42-44).

Reviewer #3:

Dillon et al. present their findings on a clinical trial investigating the safety and tolerability of concurrent ATR inhibitor (here ceralasertib) with palliative RT therapy. This clinical trial separately investigated ceralasertib as monotherapy, and the authors report here Part C of the study, which was further divided into 2 study arms according to whether the site of irradiation was above or below the diaphragm. There were 3 dose levels for ceralasertib and 2 dose levels for RT.

This is a very interesting study, investigating a hypothesized interaction between DDR, which has drugs employed largely as systemic therapy currently in combination with a locoregional therapy – palliative RT.

Further discussion on several points would improve this manuscript.

Page 4 Line 11 and 12 refer to evidence of radiosensitization in the AE profile of normal tissues as justification for the highest DL of ceralasertib. Given this important decision, perhaps Figure 1 should be referred to and briefly expanded upon.

We thank the reviewer for this important point. We have refined the text in relation to normal tissue radiosensitisation in page 3, line 53 and page 4, line 1-5. We have added text at page 4, line 12-13 referring to the mechanistic figure 1A, B:

“The exacerbation of radiation-induced normal tissue toxicity in one participant indicates that ATR also plays a role in normal tissue responses to DNA damage, possibly by reducing DNA repair or increasing DNA damage-associated replication stress (Fig. 1A, B).”

Page 5 Lines 5 refers to tumor growth kinetics before entering the study. Where these all in the context of active therapy? In the corresponding figure, much of these tumors appear to have been minimally increasing, were these the that were progressing? Or was it the out-of-field lesions? A little more background on the analysis underlying this claim would be helpful.

Thank you for this insightful observation. There was not a requirement for disease progression prior to study entry, but there was a requirement that the patient had to have a clinical indication for radiotherapy, which was usually progressive disease and risk of symptoms, symptoms requiring palliation, or attempt for local disease control (to prevent symptoms arising). We do not have granular detail on the active therapies prior to starting the study treatment. We have aggregate data that shows a number of patients had progressed on prior lines of therapy. For those with available data plotted in the supplementary figure, the majority were progressing before therapy (irradiated lesions) and the majority reduced in growth rate after treatment. We have plotted the growth rates in % per day and added these to supplementary Fig. 10C.

Page 5 Line 18, what separate malignancy? Is there a concern for a distal safety signal here?

Lung cancer in heavy smoker that was not judged to be in any way related to the investigational medicinal product; this has been appended (page 5, line 15)

TP53 is mentioned as a gene of interest as per Page 6 Line 12. The supp GSEA figure does not appear to have TP53 pathway upregulated in it though. KRAS is also discordant. Are these two truly analogous?

We thank the reviewer for the comment about TP53. This is an error and has been corrected to reflect the downregulation of G2/M signalling in both models. It is unlikely that the model is completely analogous to the human tumours, the main point of this section is to compare the immune cell depletion at early time points and suggest that stromal factors may also be an area for further study.

Page 7 line 4 mentions the durable responses are greater than would be expected with palliative radiotherapy alone. But then in line 18-20 there is mention that there is a lack of proper reference baseline studies. These appear to be contradictory claims. More discussion would be helpful or otherwise avoiding a direct comparison.

We accept this comment and have tried to clarify the matter. By its very nature, palliative radiotherapy is often given as part of end-of-life care without an expectation of prolonging survival. Based on our own experience and the medical literature, there is absolutely no precedence for T4 HNSCC being cured with 20 Gy in 10 f. We have attempted to moderate our claim in the text to be more objective, but believe that this case represents an important observation that could potentially point the way towards improved outcomes for patients treated with RT plus ceralasertib.

Page 6, line 51-52:

“This dose effectively modulates ATR signaling and is associated with durable responses, including cure of an HNSCC with 20 Gy in 10 fractions.”

The changes in localized immune responses with temporal data is very interesting and as the authors point out worth additional studies. The hypothesis as pointed out in page 8 lines 4-6 is interesting, but perhaps a little strongly stated at current.

Thank you for this comment; based on the additional references recommended by reviewer 1, we believe that evidence from other work significantly backs up this hypothesis, and we have re-worded this section to reflect this and to present the data in a more balanced fashion. At page 7, line 54-55 and page 8 line 1:

“it is reasonable to hypothesize, based on other pre-clinical and clinical data showing ATRi effect on immune populations²³, as well as the data from our study that the ATRi-RT combination may deplete resident anergic/exhausted or suppressive populations and make space for repopulation with new, vigorous T-cells”

Overall, the discussion and analysis appear to collapse the 2 study arms above/below the diaphragm into one study group to report results. It would be helpful to clarify throughout the manuscript or otherwise comment on why the study arms are not discussed or analyzed separately.

We have added a line in page 2, line 35-36 to explain why we have analysed the two arms together:

“The data from both arms have been presented together, due to no clear site-specific DLT and poor recruitment below the diaphragm”

POINT TO POINT RESPONSE TO REVIEWER - 5

Comment - Please respond to the remaining points raised by the newly invited R5 (biostats, replacing R2), i.e. previous comment 3 remains to be answered, by providing 95% CIs for CR rates.

Response - These have now been calculated and included for all response categories: page 3, line 35 – 39.