nature biomedical engineering

PEER-REVIEW INFORMATION

Prospective testing of clinical Cerenkov luminescence imaging against standard-of-care nuclear imaging for tumour location

Corresponding author: Jan Grimm

Editorial note

This document includes relevant written communications between the manuscript's corresponding author and the editor and reviewers of the manuscript during peer review. It includes decision letters relaying any editorial points and peer-review reports, and the authors' replies to these (under 'Rebuttal' headings). The editorial decisions are signed by the manuscript's handling editor, yet the editorial team and ultimately the journal's Chief Editor share responsibility for all decisions.

Any relevant documents attached to the decision letters are referred to as **Appendix #**, and can be found appended to this document. Any information deemed confidential has been redacted or removed. Earlier versions of the manuscript are not published, yet the originally submitted version may be available as a preprint. Because of editorial edits and changes during peer review, the published title of the paper and the title mentioned in below correspondence may differ.

Correspondence

Tue 09 Mar 2021
Decision on Article nBME-21-0353

Dear Prof Grimm,

Thank you again for submitting to *Nature Biomedical Engineering* your manuscript, "Clinical Cerenkov Imaging in Oncology". The manuscript has been seen by 3 experts, whose reports you will find at the end of this message. You will see that the reviewers have good words for the work, and that they raise a number of technical criticisms that we hope you will be able to address. In particular, we would expect that a revised version of the manuscript provides:

* a revision of the methodology used, taking into account the technical considerations and suggestions for improvement raised by the reviewers,

* improvements to the quantitative assessment of the data, including refining the emission wavelength filtering step,

* improvements to the data presentation and to the figures, in order to better showcase how Cherenkov imaging compares to the standard of care.

* Please remove the section describing the business model.

* Please improve the data on Xofigo treated patients, we suggest not to remove them from the paper, however, if the analysis of this dataset remains inconclusive, please move it to the SI.

When you are ready to resubmit your manuscript, please upload the revised files, a point-by-point rebuttal to the comments from all reviewers, the (revised, if needed) reporting summary, and a cover letter that explains the main improvements included in the revision and responds to any points highlighted in this decision.

Please follow the following recommendations:

* Clearly highlight any amendments to the text and figures to help the reviewers and editors find and understand the changes (yet keep in mind that excessive marking can hinder readability).

* If you and your co-authors disagree with a criticism, provide the arguments to the reviewer (optionally, indicate the relevant points in the cover letter).

* If a criticism or suggestion is not addressed, please indicate so in the rebuttal to the reviewer comments and explain the reason(s).

* Consider including responses to any criticisms raised by more than one reviewer at the beginning of the rebuttal, in a section addressed to all reviewers.

* The rebuttal should include the reviewer comments in point-by-point format (please note that we provide all reviewers will the reports as they appear at the end of this message).

* Provide the rebuttal to the reviewer comments and the cover letter as separate files.

We hope that you will be able to resubmit the manuscript within 25 weeks from the receipt of this message. If this is the case, you will be protected against potential scooping. Otherwise, we will be happy to consider a revised manuscript as long as the significance of the work is not compromised by work published elsewhere or accepted for publication at *Nature Biomedical Engineering*. Because of the COVID-19 pandemic, should you be unable to carry out experimental work in the near future we advise that you reply to this message with a revision plan in the form of a preliminary point-by-point rebuttal to the comments from all reviewers that also includes a response to any points highlighted in this decision. We should then be able to provide you with additional feedback.

We hope that you will find the referee reports helpful when revising the work, which we look forward to receive. Please do not hesitate to contact me should you have any questions.

Best wishes,

Rosy

__
Dr Rosy Favicchio
Senior Editor, *Nature Biomedical Engineering*

_____

Reviewer #1 (Report for the authors (Required)):

This paper is novel and presents a potentially clinically useful imaging option for radiopharmaceuticals. The authors have developed an imaging device that is able to detect a number of different diagnostic and therapeutic radiotracers, showing its versatility.

Minor comments:

1. Title:

Line 1: Should the title be "Clinical Cerenkov luminescence imaging in oncology" since the imaging is referred to as CLI throughout the manuscript?

2. Introduction: The authors make a strong argument that there is a need for inexpensive oncologic imaging in low-income countries. However, it is unclear if the imaging device is the only limiting factor. They state that "radiopharaceuticals can be shipped or produced locally in a generator". Could the authors give proof that

this is reasonable, i.e. are generators in existence in low-income countries? Or is the price of buying shipped radiopharmaceuticals reasonable in low-income countries?

Line 79 Survivorship imaging (used twice, line 351): perhaps disease surveillance is a better term. This reviewer has not heard of survivorship imaging commonly used in oncology.

3. Clinical imaging: The authors should describe the patient positioning in the dark box more specifically. The diagram shows a chair with patient facing the fiberscope. However, clearly some patients are positioned AP and others in the PA position. How is this achieved? Note: this become important for comments in #7.

4. Image feasibility and efficacy: The authors report that two tracers had significant correlation between CL intensity and quantified activity delivered to the activity. Have the authors considered if this could be improved with corrections for tissue properties? Perhaps this is beyond the scope of this paper.

Line 190-194- Xofigo imaging: The authors report difficulty with this imaging due to physiologic uptake among other concerns. Would imaging at a later time point improve the specificity of this imaging? The concern mentioned about uptake degenerative bone disease is also an issue with other tracers, but can typically be discriminated clinically with other imaging features, without necessitating a biopsy.

5. Spectral Cerenkov Luminescence imaging: The authors use a filtering technique to investigate depth of lesions. Did they consider orthogonal CL imaging? How would this compare?

Line 207: “Several subsequent patient images were acquired with a higher binning mode (8x8), yielding similar images in 2-5 minutes compared to 15 minutes (4x4) (Supplementary Figure 6).”
Please comment on the tradeoff for quicker acquisition time, namely a presumed decrease in sharpness and effective resolution?

Line 213: “Light with wavelengths above 650 nm (red light) undergoes less scattering and absorption in tissue and are is better at penetrating tissue” Appears to have a spelling/syntax error

Line 216: “Installation of a filter wheel with a selection of band and long-pass filters enabled the coarse spectral analysis of emitted CL <500 nm, between 500-550 nm, 550-600 nm, and >600 nm. This allowed for a rough approximation of the depth of the CL source since deeper lesions are predominantly shifted towards the red in their spectral output, versus blue-weighted sources closer to the surface.” While the claim about weighting is true, the absolute signal intensity emitted from all wavelengths should be higher for superficial lesions; i.e. the red signal from a surface lesion should be brighter than the red signal from a deeper lesion, even though the deeper lesions will have a higher relative red signal. Please comment on this, as it also relates to my question about Figure 3 image leveling.

Line 228: “However, longer red-shifted wavelengths showed the greatest CL signal in the supraclavicular region, indicating deeper lesions with the deepest ones seen with the 600 nm long-pass filter (Figure 3). Of the lesions seen by PET/CT, the filters implemented for CLI segmented superficial from deeper lesions, giving approximate depths to the CL images taken during the session.”
As presented, this is not convincing (see above comment), please provide numbers in Figure 3 to support this claim instead of relative intensity scales. Without any quantitative analysis shown, how are the authors sure that the bright regions shown on the >600nm long pass filtered images are actually deeper? Surely red light is still emitted at more superficial depths from more superficial lesions… Figure 3 should be with a global color map to compare (quantitatively) images from different filters.

Line 265:” A dashed line has been added to each graph in Figure 5 to represent the limit of detection established by the camera noise level.” I only see a solid line here, no dashed line. Also, please clarify what you mean by camera noise level – if this is in reference to electronic noise, why not darkfield subtract?

Line 305 “While CLI suffers from limitations in optics, such as lens aberrations…” Did the authors consider flatfield correction to appropriately account for variations in intensity distributions due to vignetting?

Line 316: “We have shown that by using filters to segment the continuous CL spectrum, the approximate depth of lesions can be discerned…” The results in Figure 3 aren’t really strong enough to support this, please consider rephrasing.

Line 393: “Images show superficial and deeper lesions and with the addition of filters, an approximation of lesion depth could be inferred without a CT.” The authors don’t approximate depths of any lesions based on filtered images in this paper, the only comparison is a qualitative one to PET/CT images, so this statement is really not supported by any results. If the authors want to make this claim, depths should actually be approximated using some analysis of the images and then compared to the true depths as discerned on PET/CT.

Line 452: “Images were post-processed using an automated script (macro) in ImageJ (Supplementary to remove gamma strikes and fiber scintillation.” Please clarify noise and background removal methods employed within ImageJ (i.e. special/spatial frequency filtering?), including a discussion on impacts of this choice on image resolution and fidelity. Is this a second order effect when compared with intrinsic resolution limitations such as tissue optical interactions?

7. Figure 2- The images in the figure needs more detailed labeling or additional PET slices/images for the reader to understand the images. Although every CL imaging has a red arrow that correlates with another red arrow on the corresponding conventional imaging, it isn't clear how the investigators determined these areas and what are the other signals in the imaging? Specifically-

2C- the DOTATATE PET scan shows 3 avid lesions, but the CL image shows ~5 lesions and a linear streak on the left side of the patient. Can authors clarify what these other areas are? If any are artifact, what is the source of the artifact? In general, it is difficult to compare the CLI, which is a 2D image to the PET scans, which are shown as a single axial or coronal slice in this figure. Does a MIP image from the PET scan correlate better with the CLI? Please show in a 3rd column with MIP images. Finally, the red arrow in the CLI image appears slightly right of midline, but if the red arrow on the PET scan is pointing to the large avid lesion, it is slightly left of midline. How do you know these two lesions correlate? Does this come back to the patient positioning and how accurate the CLI image overlay on the patient image?

2E- Why are the rib lesions such a different shape between the CLI image and the bone scan?

Figure 3- Again, the PET MIP imaging might be helpful to compare to the CLI.

Also, please clarify if all Cherenkov images are on the same color scale. Min to Max is ambiguous to the reader, is this global min/max or image specific?

Figure 4- Again, include the MIP imaging and also comment on why the patient positioning is different on the different days. This makes it difficult to compare from one day to the next.

Overall, the authors have presented a novel imaging of radiopharmaceuticals that is clinically feasible, which is an important step forward. However, it is difficult to correlate the CL imaging with the conventional imaging with the pictures shown. They have quantitated this with radiology reviews and Likert scoring, but the CL imaged light emissions need better explanation. If the discrepancies are difficult to discern due to inability to account for patient positioning during the course of imaging, that is ok, but the authors should have a plan to correct this in the future. Or if these are artifacts, please explain as much as possible.

Finally, it seems that one benefit of this imaging, given that it is "cheap" and easy is to do repeated measurements after therapeutic administrations. This is not often possible for conventional imaging due to the cost and time on the machines. Do the authors see a benefit to this?

Reviewer #2 (Report for the authors (Required)):

Summary: This manuscript is a novel application of Cherenkov imaging. Previous work has been performed using Cherenkov imaging of radioisotopes, but this is (to my knowledge) the most comprehensive clinical study of this technique applied to 5 targeted radioisotopes. The authors also made a number of technical improvements on their system to reduce the impact of sources of error present in earlier applications of this system. The authors correctly emphasize that this method would not be higher fidelity when compared to standard of care techniques, but could be used as a lower cost alternative to established imaging techniques. Finally, the use of wavelength filtering to establish depth is an important innovation that adds

credence to the use of this method.

Degree of Advance and Implications: I feel that this study still needs to make a stronger case for this being a clinical alternative to established radioisotope localization techniques. For instance, the premise that this could be used for 18F-FDG localization (which provided the most promising data in this study) ignores the fact that this isotope is short lived (~110 minute half life) and requires significant infrastructure to produce (cyclotron). I think it would be a very unique situation for a facility to have a cyclotron available for F-18 production but not a PET/CT scanner. Additionally, the use of a Likert scale is fairly subjective, particularly since only 4 Radiologists were consulted to establish the data provided in Table 2. I would consider this type of analysis to be supplemental to something more quantitative for a study of this scope. While information on intensity correlation is helpful, a quantitative analysis would attempt a form of image registration with the gold standard in each respective case. If this technique is to be used in a scenario in which clinical decisions are made based on the imaging information, it must be demonstrated that a minimal level of spatial information can be discerned using this technique.

Major Questions:

1) One complication with using optical fiber in this application is the production of Cherenkov light in the fiber itself. Is this what authors are referring to when they are discussing scintillation? It is likely that the background signal due to this is being handled by image processing, but I think it should be more clearly stated that this is the case.

2) How did the authors account for geometric distortions in this study? The angle of the camera relative to the patient will have a significant effect on the perceived localization of the radiotracer. Also, it is well known that Cherenkov emission has an angular dependence. Since these isotopes are emitting isotropically, it is likely that this effect washes out, but the authors should still speak to this in the manuscript.

3) It is also a known factor in the literature that skin pigment has an effect on Cherenkov emission intensity. Given that only a small number of participants had darker skin pigmentation it is likely not affecting the overall results too much but given that so few were included in this study, and only in 2 of the 5 radiotracer studies, I do not think the authors can reasonably conclude that skin pigment has no effect on Cherenkov emission in these cases.

4) The case for using CLI in Xofigo cases is compelling, but this is by far the weakest data set. You may want to consider leaving it out until you have imaged a number of patients comparable to the other data sets (20-30)

Minor Comments:

1) In the imaging feasibility and efficacy section, please provide an activity range for each isotope studied. This information is present in the abstract and Table 1, but it would also be helpful to provide it here.

2) I realize that the Likert scale is very well established, however, I think it would still be a good idea to provide a citation on its use in Radiology applications

Missing Details on Statistics: None that I can see. No issues here.

Missing Citations: The bibliography is appropriate to this study. I do not have any suggestions here.

Additional suggestions: None

Stylistic Issues: None, this is a well written manuscript.

Reviewer #3 (Report for the authors (Required)):

The team of Prof. Grimm has extended on their groundbreaking work in Cerenkov imaging and present a unique clinical proof-of-concept study with by far the highest number of patients included so far report on a novel imaging system and show images for a range of isotopes that are frequently used in the clinic. By

reporting on this data the manuscript substantially advances the scientific and translational potential of Cerenkov imaging.

Major comments:
1. While I find the reported proof-of-concept data a big step forwards for the implementation of Cerenkov imaging, I have my doubts on the valorization aspects presented. I would suggest that the authors tune down their “LightPoint Medical business plan (is the technology a new product in the company’s portfolio?)” in the introduction and discussion and rather focus on the science behind the unique data set that has been collected. This will in my view improve the manuscript further. Note: Before any claims can be made on cost benefit or replacement of PET by CLI imaging, more is needed than mere examples. For example, how do these images impact clinical decision making? I’m sure these aspects are on the radar of this expert team, but I strongly suggest they wait for this data before making claims in this direction.
2. Since the authors specifically claim novelty through presenting a novel imaging system, they should extend on their description of the engineering aspects of the system shown in figure 2. Despite the many valuable details mentioned in the supporting information, it is hard to assess the imaging system itself. In endoscopy fiberscopes are considered to suffer from reduced sensitivity. Perhaps the authors can explain in the manuscript why, apparently, this limitation is not an issue for Cerenkov imaging?
3. The authors discuss filtering of the emission wavelength and how this can impact clinical utility (Figure 3). But in current form this data seems to be suggestive more than anything. One could potentially get similar images when the sensitivity of detection is reduced independent of the wavelength. Is there a way to exclude this potential artefact and truly validate the effect is wavelength dependent? E.g. by using a rainbow visualization of the signal intensity? Please note: For fluorescence imaging evidence there is increasing emerging that fluorescent emission in the 500-600 range also provides clinical value e.g. fluorescein, suggesting that perhaps the need to shift to high wavelength regions is not perse a requirement for clinical utility.
4. It seems to me that Figure 5 provides data that could convert into new scientific insights and data of which the interpretation is instrumental for the future utilization of Cerenkov imaging. I would urge the authors to try and get more information from this data and make stronger conclusions in the results and discussion section based on this information. For example, is there potential clinical value in using the measured signal intensities (and energy of the radiation that is the source of the Cerenkov light) to determine the depth of origin? Or is it very much like clinical fluorescence imaging: you either see something, or you don’t? For the latter the lack of quantitative data (and tissue attenuation of signal) limits the clinical utility and essentially require it to be used in combination with a radiological modality. With many of the radiotracers used you can overcome this limitation through e.g. PET imaging or beta-probe based signal detection, please discuss this in more detail.

Minor comments:

1. Minor: The example images of patients that have been provided in Figures 2 and 4 are impressive. Would it be possible to include more images in the SI? Just so that a reader can better assess how image quality may vary between patients.
2. Minor: The authors should briefly mention the radiation exposure concept behind Cerenkov imaging. I agree that this is irrelevant when you image patients that were scheduled for PET, but this feature becomes relevant when you propose to use Cerenkov imaging as independent modality for diagnostics or image guided surgery. Mainly because the alternatives would yield less radiation exposure for patients.

Thu 05 Aug 2021
Decision on Article nBME-21-0353A

Dear Prof Grimm,

Thank you for your revised manuscript, "Clinical Cerenkov Imaging in Oncology", which has been seen by the original reviewers. In their reports, which you will find at the end of this message, you will see that the reviewers acknowledge the improvements to the work and raise a few additional technical criticisms that we hope you will be able to address. In particular, we would expect that the next version of the manuscript provides:

* discussion of the dependence on camera orientation relative to the patient;

* discuss the limitations in detecting Xofigo.

As before, when you are ready to resubmit your manuscript, please upload the revised files, a point-by-point rebuttal to the comments from all reviewers, the reporting summary, and a cover letter that explains the main improvements included in the revision and responds to any points highlighted in this decision.

As a reminder, please follow the following recommendations:

* Clearly highlight any amendments to the text and figures to help the reviewers and editors find and understand the changes (yet keep in mind that excessive marking can hinder readability).

* If you and your co-authors disagree with a criticism, provide the arguments to the reviewer (optionally, indicate the relevant points in the cover letter).

* If a criticism or suggestion is not addressed, please indicate so in the rebuttal to the reviewer comments and explain the reason(s).

* Consider including responses to any criticisms raised by more than one reviewer at the beginning of the rebuttal, in a section addressed to all reviewers.

* The rebuttal should include the reviewer comments in point-by-point format (please note that we provide all reviewers will the reports as they appear at the end of this message).

* Provide the rebuttal to the reviewer comments and the cover letter as separate files.

We hope that you will be able to resubmit the manuscript within 12 weeks from the receipt of this message. If this is the case, you will be protected against potential scooping. Otherwise, we will be happy to consider a revised manuscript as long as the significance of the work is not compromised by work published elsewhere or accepted for publication at *Nature Biomedical Engineering*.

We look forward to receive a further revised version of the work. Please do not hesitate to contact me should you have any questions.

Best wishes,

Rosy

Dr Rosy Favicchio
Senior Editor, *Nature Biomedical Engineering*

Reviewer #1 (Report for the authors (Required)):

The authors addressed prior concerns. Overall, this is an important work given the novelty and proof that this imaging is possible in patients and informative of disease activity. Further work will be needed for clinical use, but this is an important step in the development.

Reviewer #2 (Report for the authors (Required)):

Summary: The authors have put a good deal of effort into addressing my concerns of the original manuscript. The revised manuscript is much improved. The new version of this work now has a more pragmatic tone in regards to the utility of this method. Additionally, efforts to better quantify the sensitivity of this new method over marked improvements to this work. The authors have addressed most of my concerns, but I still have 2 major issues at this point that I feel should be addressed.

Degree of Advance and Implications: This revised manuscript does a much better job at justifying the novelty and utility of Cherenkov imaging in regions where PET/CT scanners or SPECT scanners are scarce. The authors bring up a very good point that this would be a useful tool in triage scenarios when deciding if a PET scan is needed. You could also argue that this would be a reasonable approach in any hospital system, even one in a wealthy country, and could help reduce costs to patients.

Major concerns:

1) In regards to orientation dependence, the authors supply examples in Supplementary Figure 9 of Cherenkov signal vs orientation. I would say that based on these images that there is a significant dependence on camera orientation relative to the patient. For instance, in Example 2, we can see in one case a single focal point for the Cherenkov signal while in the next view there appears to be 2 focal points. SPECT here would indicate only one focal point of uptake. Perhaps recommending a preferred camera orientation would be a good route to go here?

2) The authors state the the "limitations in Xofigo have been pointed out more clearly to address this point" in the current draft, but I do not see where that is. Can you please point to specific language that addressed this? One example is in Supplementary Figure 8E which demonstrates that the perceived uptake when comparing CLI to PET is very different. I do not see these discrepancies discussed in the manuscript.

Minor concerns:

1) Looking at Figure 5, it looks like subfigure C has a significant correlation. The P value is lower than in subfigure B. If this is true, can the authors amend their text to reflect this?

2) I was able to track down most of the changes, but there were some changes that did not correspond to the locations indicated by the authors in the Response to Reviewers document. For example, Supplementary Figure 5A-B was emphasized for the question on scintillation, but I think the correct figure to refer to here is Figure 7A-B. Also, in many places the indicated pages where changes were made were not correct. For the most part, I was able to track down changes with a search tool, but more accurate page designations would be more helpful in the future.

Statistics, protocols, or materials: As before, no concerns here.

Citations: The added citations improve this work. I do not see the need for any additional changes here.

Optional suggestions: None

Stylistic issues: As before, no issues here

Reviewer #3 (Report for the authors (Required)):

I would like to thank the authors for thoroughly addressing my comments/concerns.

Sun 13 Feb 2021
Decision on Article nBME-21-0353B

Dear Prof Grimm,

Thank you for your patience in waiting for our feedback on your revised manuscript, "Clinical Cerenkov Luminescence Imaging in Oncology". As communicated by my colleague Rosy Favicchio, having consulted with Reviewer #2 (whose brief comments you will find at the end of this message), I am pleased to write that we shall be happy to publish the manuscript in *Nature Biomedical Engineering*, provided that the points specified in the attached instructions file are addressed.

When you are ready to submit the final version of your manuscript, please upload the files specified in the instructions file.

For primary research originally submitted after December 1, 2019, we encourage authors to take up transparent peer review. If you are eligible and opt in to transparent peer review, we will publish, as a single supplementary file, all the reviewer comments for all the versions of the manuscript, your rebuttal letters, and the editorial decision letters. **If you opt in to transparent peer review, in the attached file please tick the box 'I wish to participate in transparent peer review'; if you prefer not to, please tick 'I do NOT wish to participate in transparent peer review'.** In the interest of confidentiality, we allow redactions to the rebuttal letters and to the reviewer comments. If you are concerned about the release of confidential data, please indicate what specific information you would like to have removed; we cannot incorporate redactions for any other reasons.
More information on transparent peer review is available.

Please do not hesitate to contact me should you have any questions.

Best wishes,

Pep

__
Pep Pàmies
Chief Editor, *Nature Biomedical Engineering*

_____

Reviewer #2 (Comments to the authors):

The authors have addressed all of my remaining concerns. This manuscript is an important and innovative addition to the field of Clinical Cherenkov Imaging and should be published in its current form.

_____

Rebuttal 1

Reviewer #1:

This paper is novel and presents a potentially clinically useful imaging option for radiopharmaceuticals. The authors have developed an imaging device that is able to detect a number of different diagnostic and therapeutic radiotracers, showing its versatility.

We thank the reviewer for the positive comment.

Minor comments:

1. Title:

Line 1: Should the title be "Clinical Cerenkov luminescence imaging in oncology" since the imaging is referred to as CLI throughout the manuscript?

We thank the reviewer for the suggestion and have changed the manuscript title as suggested to "Clinical Cerenkov luminescence imaging in oncology".

2. Introduction: The authors make a strong argument that there is a need for inexpensive oncologic imaging in low-income countries. However, it is unclear if the imaging device is the only limiting factor. They state that "radiopharaceuticals can be shipped or produced locally in a generator". Could the authors give proof that this is reasonable, i.e., are generators in existence in low-income countries? Or is the price of buying shipped radiopharmaceuticals reasonable in low-income countries?

We thank the reviewer for this observation. Improving access is at minimum a two-part problem; providing a scanner while also providing radiotracer supply locally or through a supply chain. In well-established facilities, a cyclotron or nearby cyclotron represents the major infrastructure investment along with several PET or SPECT scanners. Finding cheaper types of scanners (i.e. this work) as well as ways to provide isotope supply will greatly improve access where previously cost prohibitive. The Nuclear Medicine Global Initiative as well as the Lancet Oncology Commission have recently highlighted inequity in radiotracer supply as well as imaging device access. The recent Nuclear Medicine Global Initiative's report on radiopharmaceutical access (references 4 and 5) identified the growing use of ^{68}Ga Generators with 13/32 of surveyed countries using these generators to support DOTATATE and PSMA scans. In addition the longer-lived isotope Iodine-131 was used in 30/32 countries. Commentary from the Lancet Commission found that ^{68}Ga generators "might be an opportunity to implement more advanced imaging techniques at an earlier stage" given the months long life of these generators and available kits for tracer production. Papers covering the construction of cyclotron facilities in Brazil and Bangladesh highlight the existence of radioisotope supply in low-income countries, as well as the interest in generator-based systems given their lower cost compared to a full cyclotron facility. Furthermore, Brazil has also made their own ^{68}Ga generator, with the development described by Brambilla and Osso (Reference 9). While commercial and GMP sources are available, the Lancet Oncology Commission highlights the need to consider the use of non GMP sources, provided they adhere to local quality standards. Lower cost entry methods for nuclear medicine and imaging are attractive options for income limited countries. We have amended the introduction to focus more on improving access to imaging with varied isotope sources.

Line 79 Survivorship imaging (used twice, line 351): perhaps disease surveillance is a better term. This reviewer has not heard of survivorship imaging commonly used in oncology.

We thank the reviewer for pointing out this error and we have updated the manuscript to disease surveillance instead of “survivorship”.

3. Clinical imaging: The authors should describe the patient positioning in the dark box more specifically. The diagram shows a chair with patient facing the fiberscope. However, clearly some patients are positioned AP and others in the PA position. How is this achieved? Note: this become important for comments in #7.

We thank the reviewer for this question and the cross reference as well. Positioning is very important for optical imaging, and to minimize differences between images, imaging was limited to known lesion locations, which were placed near the center of view at 1.4’-2’ from the fiberscope lens. Space limitations in the enclosure and designated room allowed only for a chair with a semi reclining back rest. Whilst sometimes suboptimal, this meant that patients could be positioned in AP or PA position on the chair. Whilst outside the current scope, a future implementation of the device could incorporate a bed with a fixed imaging distance. However, this could prove difficult due to demand for space in clinical settings in larger cities.

4. Image feasibility and efficacy: The authors report that two tracers had significant correlation between CL intensity and quantified activity delivered to the activity. Have the authors considered if this could be improved with corrections for tissue properties? Perhaps this is beyond the scope of this paper.

We thank the author for this comment. We were able to reanalyze the iodine patient set, extracting the counts from the original planar SPECT images, instead of relying on an estimated product of Iodine 123. Uptake and the mCi of Iodine 131 administered. As such we were able to show significance for Figure 5C where increasing CLI did correlate with increasing iodine present in tissue. The three radiotracers, ^{18}F-FDG Iodine-131, and ^{177}Lu-DOTATATE, which demonstrated significant correlation, could be improved upon via specific corrections for tissue properties. The remaining two isotopes may be improved, but CLI adjustment based on the exponential absorption with distance did not improve the ^{68}Ga-DOTATATE correlation for example (not shown). Radium-223 scans are too preliminary with the limited number of patients with a ^{18}F PET/CT currently. Recently published work by Pogue et al. has shown the improvement in Cerenkov dosimetry using CT weighting factors for tissue attenuation correction. Unfortunately, not all patients received CT scans prior to CLI and thus this method is not applicable here without a generalized correction. Additionally, the variability from patient to patient (size, positioning, body habitus), tumor depth and location would have to be considered. Additionally, the model would likely need to be weighted for each body part e.g. arms, legs, torso etc. To generate an accurate model without this knowledge is indeed beyond the scope of this work. However, we have added this to the discussion as potential next steps and updated the references accordingly. This future analysis would greatly enhance the imaging modality when a commercial camera is established.

Line 190-194- Xofigo imaging: The authors report difficulty with this imaging due to physiologic uptake among other concerns. Would imaging at a later time point improve the specificity of this imaging? The concern mentioned about uptake in degenerative bone disease is also an issue with other tracers, but can typically be discriminated clinically with other imaging features, without necessitating a biopsy.

The reviewer is correct that imaging at later timepoints to reduce the signal from physiologic distribution should help improve pathologic imaging. We have added a recent publication by Flux et al. highlighting the challenge of Radium imaging due to the several days of clearance of the tracer through

the gastrointestinal tract (pages 6 and 31) suggesting later imaging may help, but the optimal time may be the following day. Our study was entirely dependent on the goodwill and interest of the patients to give us their precious time, and patients receiving Xofigo have mostly not been willing to stay much longer than their infusion of 15 minutes. They often did not have any follow up visits scheduled in the hospital in the following days, which would allow for a facile later imaging timepoint. The IRB protocol had not allowed us to ask patients to come back only for the purpose of this study.

5. Spectral Cerenkov Luminescence imaging: The authors use a filtering technique to investigate depth of lesions. Did they consider orthogonal CL imaging? How would this compare?

We thank the author for this suggestion. The time in which patients could be imaged was limited and given to us voluntarily by the study participants, thus imaging a set of 5 filter images (minimum of 25 mins) along with an additional orthogonal image was not feasible or acceptable for most patients. However, we have provided two additional patients that were imaged with all of these filters (Supplementary Figure 12) as well as iodine patients imaged from two different angles (Supplementary Figure 9). We also conducted a tissue phantom with the filters (Supplementary Figure 11) and ultimately determined that filters below 600nm to not provide any meaningful information as to lesion depth and as such have updated the claim in the manuscript on pages 11 and 17. Due to the limits of light penetration in tissue in the visible region, we observed the majority of the permeating light coming from the surface to the camera to be longer than 600nm. Here, planar imaging with a recent CT could greatly provide the depth information and inform how potentially avid/bright a particular lesion may be.

Line 207: “Several subsequent patient images were acquired with a higher binning mode (8x8), yielding similar images in 2-5 minutes compared to 15 minutes (4x4) (Supplementary Figure 6).”
Please comment on the tradeoff for quicker acquisition time, namely a presumed decrease in sharpness and effective resolution?

CLI is an optical based method and in this example being performed on a very large scale in comparison to pre-clinical CLI (cm vs mm). We believe CLI in clinical settings is positioned as a detection method where signal is either present or absent. The inherent scattering of visible light further reduces the effective resolution in a clinical setting and as shown, displays the approximate location of the lesions. We have found that there is little to no tradeoff when switching between 4x4 and 8x8 binning modes due to the image distance and the processing steps applied and the effective increase in sensitivity is worth any negligible loss in resolution at this scale. It should also be noted that the camera in this case performs physical binning as opposed to software binning. It cannot be guaranteed that similar results can be achieved with software binning. We have pointed out that a change in binning allows for a shorter acquisition with a nominal loss in resolution due to post processing steps, on page 10 and as standalone supplementary figures 5 and 10.

Line 213: “Light with wavelengths above 650 nm (red light) undergoes less scattering and absorption in tissue and are is better at penetrating tissue” Appears to have a spelling/syntax error

Corrected, thank you.

Line 216: “Installation of a filter wheel with a selection of band and long-pass filters enabled the coarse spectral analysis of emitted CL <500 nm, between 500-550 nm, 550-600 nm, and >600 nm. This allowed for a rough approximation of the depth of the CL source since deeper lesions are predominantly shifted towards the red in their spectral output, versus blue-weighted sources closer to the surface.” While the

claim about weighting is true, the absolute signal intensity emitted from all wavelengths should be higher for superficial lesions, i.e. the red signal from a surface lesion should be brighter than the red signal from a deeper lesion, even though the deeper lesions will have a higher relative red signal. Please comment on this, as it also relates to my question about Figure 3 image leveling.

We thank the reviewer for highlighting this. We conducted a filter experiment with a tissue phantom containing lesions of different depth, containing the same activity. Briefly we found that the 600nm filter best reproduced the expected CLI of the filters tested. Our filter analysis of Cerenkov lesions utilized the standard of care SPECT/CT or PET/CT to determine the activity present in, as well as the depth of each lesion and correlated these to the Cerenkov image acquired after that standard of care scan. It is true that two identically sized and avid lesions of different depth would recapitulate the scenario described above. The human sized abdominal tissue phantom with five different lesions of identical size and activity but at varying depths was added as Supplementary Figure 11. Indeed, we found the shallowest spheres to be the brightest, with the 600nm filter performing best of all the tested filters and most accurately reproduced the spheres location. We did not find the short-pass nor the bandpass filters below 600nm to provide accurate depth information as the majority of CL emitted from tissue (and phantom) is above 600nm. The phantom we have imaged consisting of different depths exemplifies the complexity of a Cerenkov only depth assessment based on intensity alone and should be avoided without having an orthogonal activity measurement or image understanding approximate depth, similar to a diagnostic X-ray approach, which is also done mostly in two planes.

Line 228: “However, longer red-shifted wavelengths showed the greatest CL signal in the supraclavicular region, indicating deeper lesions with the deepest ones seen with the 600 nm long-pass filter (Figure 3). Of the lesions seen by PET/CT, the filters implemented for CLI segmented superficial from deeper lesions, giving approximate depths to the CL images taken during the session.”
As presented, this is not convincing (see above comment), please provide numbers in Figure 3 to support this claim instead of relative intensity scales. Without any quantitative analysis shown, how are the authors sure that the bright regions shown on the >600nm long pass filtered images are actually deeper? Surely red light is still emitted at more superficial depths from more superficial lesions… Figure 3 should be with a global color map to compare (quantitatively) images from different filters.

We thank the reviewer for the clarification, we have included the intensities for each CL image from figures 2 through 5 as well in the supplement. For the filter analysis in figure 3 and supported with supplementary figures 11 and 12, we examined the emission of CLI from patients in bands below 600nm and above 600nm. We found the 600nm filter provided the best image of lesion location and potential depth for both a tissue phantom and in patient images. As the majority of CLI emitted is red shifted, CLI filtering below 600nm is not feasible, and longer wavelength filtering should be explored. In figure 3 we included the SUV and lesion depth, highlighting that the main lesions found were primarily lesion 1 and 3, similarly avid with a SUV of ~50. With the 600nm filter, the deeper of the two lesions was less intense, but confirmation of this observation required the combination of the PET SUV and the CT depth. We have revised our manuscript to reflect that in the images required, the majority of light emitted is above 600nm, in agreement with known optical properties of tissue, and have added references for this in both the filter paragraph on page 11 and in the discussion on page 15.

Line 265:” A dashed line has been added to each graph in Figure 5 to represent the limit of detection established by the camera noise level.” I only see a solid line here, no dashed line. Also, please clarify what you mean by camera noise level – if this is in reference to electronic noise, why not darkfield subtract?

We have updated the line accordingly and we thank the reviewer for highlighting this. This is in reference to electrical and read noise. The camera used (Andor 889) is state of the art however, over these long acquisition times it is expected that some noise will be registered during acquisitions. We have clarified this sentence and added a section using a calibrated light source to clarify the noise level. Due to the inherent random effect of gamma strikes on the sensor, we did not perform darkfield subtraction. However, the processing steps employed including median filtering and bandpass FFT filtering provide sufficient sensitivity to detect CL from a variety of patients and isotopes.

Line 305 "While CLI suffers from limitations in optics, such as lens aberrations…" Did the authors consider flatfield correction to appropriately account for variations in intensity distributions due to vignetting?

We thank the reviewer for these comments. Our comment was intended for general issues with optical based systems. We already had implemented processing steps in the macro to account for vignetting in the system, namely FFT bandpass filtering to correct for horizontal bands which are present in the camera along with uneven shading or illumination effects which this plugin was designed to remove. This issue will be different for each camera and future replications will likely require the optimization of processing code. We thank the reviewer for highlighting this and have added the following sentence to clarify. "Furthermore, Fast Fourier Transform (FFT) image processing of the CLI images reduces aberrations and vignetting imparted by the given system setup" on page 14.

Line 316: "We have shown that by using filters to segment the continuous CL spectrum, the approximate depth of lesions can be discerned…" The results in Figure 3 aren't really strong enough to support this, please consider rephrasing.

We apologize for this overstatement. We have included supplementary figure 11 showing that the shorter wavelength filters were insufficient to resolve the depth of lesions and have adjusted the manuscript to mention the majority of CL emitted is above 600nm. As such Figure 3 has been truncated to show the open filter and 600nm band pass, and how with the PET/CT more superficial lesions are clearly brighter in the 600nm image. The additional filters, along with two other patient filter examples have been added as Supplementary Figure 12

Line 393: "Images show superficial and deeper lesions and with the addition of filters, an approximation of lesion depth could be inferred without a CT." The authors don't approximate depths of any lesions based on filtered images in this paper, the only comparison is a qualitative one to PET/CT images, so this statement is really not supported by any results. If the authors want to make this claim, depths should actually be approximated using some analysis of the images and then compared to the true depths as discerned on PET/CT.

We thank the reviewer for this observation. We created in Supplementary Figure 11 a phantom with different tumor depths in a controlled situation and measured filter performance with depth. We found that the CLI intensity correlated with lesion depth only in the 600nm filter. However, for any meaningful depth information, CLI requires depth measurement or activity in the lesion from another imaging modality. As such we have revised the manuscript to reflect this information and truncate figure 3 to show the open and 600nm filter images alongside the PET MIP of the patient with measured SUV and depth by CT.

Line 452: “Images were post-processed using an automated script (macro) in ImageJ (Supplementary to remove gamma strikes and fiber scintillation.” Please clarify noise and background removal methods employed within ImageJ (i.e. special/spatial frequency filtering?), including a discussion on impacts of this choice on image resolution and fidelity. Is this a second order effect when compared with intrinsic resolution limitations such as tissue optical interactions?

We thank the reviewer for this question and have expanded our image processing description in the methods section as well as in supplementary figure 15. In summary the following steps are used in order to render meaningful CL images. Gamma strikes are removed via the built-in outlier removal in ImageJ which runs a median filter to suppress bright pixels. With the current setup we found a threshold of 3000 with a radius size of 8 effectively removed remaining gamma strikes that lead shielding and the baffle failed to prevent. In order to remove background, noise and vignetting (shading) effects we employed the native bandpass FFT in ImageJ, now referenced in the paper. The filter acts via gaussian filtering in the Fourier domain to remove uneven shading or lighting (vignetting). Additionally, horizontal stripes from fiber scintillation were suppressed with a tolerance of 10 in order to remove sensor deformations present in our system and the function suppresses Fourier components on the horizontal Fourier coordinate axis. Due to the small size of the CL images (128x128 or 64x64 pixels) the FFT bandpass filtering could be quickly completed, and processing of multiple patient datasets was achieved in a few seconds. Vignetting effects most prominent at the edge of the imaging field, were further removed via border cropping, 4 pixels from the edge in the case of 64-pixel images and 8 pixels for 128-pixel images. Finally, by automatically thresholding the data to the top 5% of the gray values any remaining noise is removed from the final image. From our experience with this clinical trial the main limitation in detecting CLI is sensitivity and thus the added increase in sensitivity achieved via binning far outweighs the reduction in resolution of the system. Fortunately, our system performs on sensor binning as opposed to software binning creating “superpixels” that are highly sensitive with very low noise levels. As the reviewer mentions, the other compounding factor determining our resolution is the inherent scattering of light by tissue. Due to the mesoscale nature of the system, we have found the employed processing steps do not have a significant effect on the resolution and in fact aid in improving the image fidelity with the resulting CL being correctly localized to avid tumor regions. Additionally, the processing steps used here have been used in similar manners in other preclinical CL studies.

7. Figure 2- The images in the figure needs more detailed labeling or additional PET slices/images for the reader to understand the images. Although every CL imaging has a red arrow that correlates with another red arrow on the corresponding conventional imaging, it isn't clear how the investigators determined these areas and what are the other signals in the imaging? Specifically-

We thank the reviewer for these comments. We have updated figures 2, 3, and 4 to include specific intensities for both CLI and the standard of care images. Where available we have replaced PET slices with maximum intensity projections and complete planar SPECT images to show all possible light sources and described where the CLI may differ from the standard of care image.

2C- the DOTATATE PET scan shows 3 avid lesions, but the CL image shows ~5 lesions and a linear streak on the left side of the patient. Can authors clarify what these other areas are? If any are artifact, what is the source of the artifact? In general, it is difficult to compare the CLI, which is a 2D image to the PET scans, which are shown as a single axial or coronal slice in this figure. Does a MIP image from the PET scan correlate better with the CLI? Please show in a 3rd column with MIP images. Finally, the red arrow in the CLI image appears slightly right of midline, but if the red arrow on the PET scan is pointing to the

large avid lesion, it is slightly left of midline. How do you know these two lesions correlate? Does this come back to the patient positioning and how accurate the CLI image overlay on the patient image?

We thank the reviewer for these questions. We have replaced the suggestive slice images with MIPs to show the total possible light from a given region. For 2C, the MIP better identifies that the linear streak of lesions is similarly bright, and that scattering in tissue across the lesion spreads the light seen away from the center of the lesion, towards the chest and down into the abdomen. Additional factors such as patient movement could also explain the possible artifact light.

2E- Why are the rib lesions such a different shape between the CLI image and the bone scan?

We thank the reviewer for this observation and have added in the figure legend that the presence of the scapula and the rhomboid muscle likely obscures the CLI coming from this source. Also, the lesions are predominantly in the marrow cavity and have to pass through the rib's cortex, which leads to further scatter and absorption.

Figure 3- Again, the PET MIP imaging might be helpful to compare to the CLI.

We apologize for the confusion in imaging. Typically quantitative slices are represented in PET/CT images, but as we are comparing a planar CLI image we have included the MIPs where available to show all possible light sources. The image threshold range was established using a 95% of the maximum intensity as defined in the macro for all images. Artifacts that do not visually match the activity present in the standard of care image can result from the fixed threshold percentage, undetermined scatter of light, surface contamination of the radiotracer by the patient (mainly iodine patients), and the incomplete removal of gamma strikes from the image during image processing. We standardized the image processing to minimize inter image bias but acknowledge that the macro as applied may contain undetermined systemic biases. In addition to the final overlay, the non-thresholded images are also exported where the user can further manually adjust the thresholding in a similar fashion as to how a radiologist adjusts PET windowing levels.

Also, please clarify if all Cherenkov images are on the same color scale. Min to Max is ambiguous to the reader, is this global min/max or image specific?

We apologize for this confusion and we have amended the images with their individual intensity scales. As each patient is unique, each color scale has a different range (similar to a PET scan) and is specified in the respective figure and represents the 16-bit gray values in the image. CLI images in Figure 4 are matched to the prior gallium or lutetium scan intensity for comparison.

Figure 4- Again, include the MIP imaging and also comment on why the patient positioning is different on the different days. This makes it difficult to compare from one day to the next.

We apologize for the confusion in figure 4. It was not always possible to get the patient to sit and hold still in the same position as previously imaged, as well as operator error in matching the image position to the previous CLI image. We have mentioned these shortcomings in the figure legend and have included this in the discussion on page 15. Going forward more fixed angle imaging is proposed for clinical CLI to reduce this variation and align with current radiology best practices.

Overall, the authors have presented a novel imaging of radiopharmaceuticals that is clinically feasible,

which is an important step forward. However, it is difficult to correlate the CL imaging with the conventional imaging with the pictures shown. They have quantitated this with radiology reviews and Likert scoring, but the CL imaged light emissions need better explanation. If the discrepancies are difficult to discern due to inability to account for patient positioning during the course of imaging, that is ok, but the authors should have a plan to correct this in the future. Or if these are artifacts, please explain as much as possible.

We thank the reviewer for these comments and have added references to the utility of the Likert ranking score in radiology. As above, we have implemented MIPs where available to show the total activity present rather than single slices. We have commented on CL present in the images and how positioning, processing, or artifacts likely contribute to the final image. In addition a better placement of the patient, preventing potential movement would also reduce the artifacts seen.

Finally, it seems that one benefit of this imaging, given that it is "cheap" and easy is to do repeated measurements after therapeutic administrations. This is not often possible for conventional imaging due to the cost and time on the machines. Do the authors see a benefit to this?

Indeed, and we thank the reviewer for this thoughtful insight. We see several advantages to CLI as an imaging modality based on cost, speed and versatility and have added a sentence in the conclusion on page 15 to highlight the advantage during therapy monitoring.

Reviewer #2:

Summary: This manuscript is a novel application of Cherenkov imaging. Previous work has been performed using Cherenkov imaging of radioisotopes, but this is (to my knowledge) the most comprehensive clinical study of this technique applied to 5 targeted radioisotopes. The authors also made a number of technical improvements on their system to reduce the impact of sources of error present in earlier applications of this system. The authors correctly emphasize that this method would not be higher fidelity when compared to standard of care techniques, but could be used as a lower cost alternative to established imaging techniques. Finally, the use of wavelength filtering to establish depth is an important innovation that adds credence to the use of this method.

We like to thank the reviewer for the positive comment and recognition of the novelty.

Degree of Advance and Implications: I feel that this study still needs to make a stronger case for this being a clinical alternative to established radioisotope localization techniques. For instance, the premise that this could be used for 18F-FDG localization (which provided the most promising data in this study) ignores the fact that this isotope is short lived (~110 minute half life) and requires significant infrastructure to produce (cyclotron). I think it would be a very unique situation for a facility to have a cyclotron available for F-18 production but not a PET/CT scanner. Additionally, the use of a Likert scale is fairly subjective, particularly since only 4 Radiologists were consulted to establish the data provided in Table 2. I would consider this type of analysis to be supplemental to something more quantitative for a study of this scope. While information on intensity correlation is helpful, a quantitative analysis would attempt a form of image registration with the gold standard in each respective case. If this technique is to be used in a scenario in which clinical decisions are made based on the imaging information, it must be demonstrated that a minimal level of spatial information can be discerned using this technique.

We thank the reviewer for the comments and agree with the overall sentiment. An isotope supply chain is certainly designed with the appropriate scanners in mind. We have therefore amended the discussion to bolster the case for using CLI as a lower cost molecular imaging device, allowing multiple cameras to be deployed for the same cost as a single PET or SPECT system. Therefore, imaging capacity of a center with a cyclotron and PET/CT system could be expanded with CLI. However, we see the main use of CCLI in areas that could be supplied with FDG but only have a few PET scanners available for the triage of patients. This is especially relevant in countries where PET scanners are a rare commodity and CCLI could offer lower cost nuclear imaging, e.g. in India with 0.08 scanners / 10^6 people (PMID 27833308) or $0.3/10^6$ in Latin America versus $7/10^6$ in the USA (PMID 26229143) . Therefore, our intent is that CLI would provide a rudimentary image such that the clinician can decide if following up with a PET scan is necessary. The versatility of CLI with clinical isotopes also allows flexibility in setups, so a dedicated cyclotron facility is not a requirement. While the Likert scale is subjective, it is an established and accepted main method by radiologists for image comparison. We have added citations around the use of the Likert scale in the introduction on page 7. To enhance the quantitative analysis of this study we have added supplementary figures 4,5, and 11 covering the limit of detection down to 0.03pW, spatial resolution with 4x4 and 8x8 binning at 400um and 1mm, respectively, and a phantom re-examining tumor depth with the installed filters. We have also reexamined the intensity analysis for patients receiving Iodine-131, and extracted and converted the raw imaging counts to yield a significant intensity correlation with increasing Iodine-131 present in the thyroid bed. Together these figures better define the performance of this CLI device, and where CLI can expand imaging capacity in a clinical setting.

Major Questions:

1) One complication with using optical fiber in this application is the production of Cherenkov light in the fiber itself. Is this what authors are referring to when they are discussing scintillation? It is likely that the background signal due to this is being handled by image processing, but I think it should be more clearly stated that this is the case.

The reviewer is correct, and we apologize if we have not explained that more clearly. The light production in the fiber was found to occur when activity was placed near to the fiber with the lens cap covered (Supplementary figure 5A-B) which could occur from positrons striking the fiber, but more so the gammas also emitted in the decay or annihilation process; this could also originate in the patients. As such we implemented outlier removal (median filtering with a threshold of 3000 and radius of 8) for gamma strike removal, an FFT band pass image processing step to correct for vignetting, horizontal bands present in the system and background suppression. Finally, we performed border cropping across all images to remove the potentially most prominent vignetting effects at the edges of the imaging field. We have now further clarified this in the manuscript and have mentioned in the introduction (on page 5) that the image processing further removed this signal after the addition of the lead shielding to the camera and fiber. We have amended the supplementary Figure 7 description to state the FFT processing removes this effect.

2) How did the authors account for geometric distortions in this study? The angle of the camera relative to the patient will have a significant effect on the perceived localization of the radiotracer. Also, it is well known that Cherenkov emission has an angular dependence. Since these isotopes are emitting isotropically, it is likely that this effect washes out, but the authors should still speak to this in the manuscript.

We thank the reviewer for bringing up the angular nature of Cerenkov light. However, this is relevant when observing a single decay but in total the light is, as the reviewer points out, emitted isotropically. Also, Zhang, R. et al did show in Cerenkov phantoms the angular distribution is insensitive to the incident angle, likely due to the high scattering of light in tissue. We have included this reference in the introduction to further describe the CL signal seen. We have also included additional CL images taken at different angles of patients, showing how the CL signal does not significantly change with the patient orientation in Supplementary Figure 9.

3) It is also a known factor in the literature that skin pigment has an effect on Cherenkov emission intensity. Given that only a small number of participants had darker skin pigmentation it is likely not affecting the overall results too much but given that so few were included in this study, and only in 2 of the 5 radiotracer studies, I do not think the authors can reasonably conclude that skin pigment has no effect on Cherenkov emission in these cases.

We thank the reviewer for this observation and agree the percentage of darker pigmented patients is insufficiently powered to tell if there is any benefit or weakness for CLI. We have therefore toned down the statement on page 32 in the figure 5 caption to reflect this and acknowledge this limitation in the study accrual on page 16.

4) The case for using CLI in Xofigo cases is compelling, but this is by far the weakest data set. You may want to consider leaving it out until you have imaged a number of patients comparable to the other data sets (20-30)

We agree the Xofigo is to date the weakest dataset and has been throughout the recruitment process. No new Xofigo patients have been recruited to date due to the COVID-19 pandemic. To prevent delay, we feel the Xofigo images should be kept in just to signify the breadth of CLI, while additional examples have been added in the Supplementary Figure 8 and the intensity analysis moved to the Supplementary Figure 14. The limitations in Xofigo imaging have been also pointed out more clearly to address this point.

Minor Comments:

1) In the imaging feasibility and efficacy section, please provide an activity range for each isotope studied. This information is present in the abstract and Table 1, but it would also be helpful to provide it here.

We thank the reviewer for the suggestion and have added the administered activity ranges for each isotope given the variation in the imaging feasibility and efficacy section on page 7.

2) I realize that the Likert scale is very well established, however, I think it would still be a good idea to provide a citation on its use in Radiology applications

We thank the reviewer for this omission and have added the four references listed below on the use and benefit of Likert scaling in radiology applications.

1. Costa, D. N. et al. Diagnostic Utility of a Likert Scale Versus Qualitative Descriptors and Length of Capsular Contact for Determining Extraprostatic Tumor Extension at Multiparametric Prostate MRI. AJR Am J Roentgenol 210, 1066-1072, doi:10.2214/AJR.17.18849 (2018).

2. Koksel, Y. et al. Utility of Likert scale (Deauville criteria) in assessment of Chemoradiotherapy response of primary oropharyngeal squamous cell Cancer site. Clinical Imaging 55, 89-94, doi:https://doi.org/10.1016/j.clinimag.2019.01.007 (2019).
3. Phelps, A. S. et al. Pairwise comparison versus Likert scale for biomedical image assessment. AJR Am J Roentgenol 204, 8-14, doi:10.2214/AJR.14.13022 (2015).
4. Rosenkrantz, A. B. et al. Prostate Cancer Localization Using Multiparametric MR Imaging: Comparison of Prostate Imaging Reporting and Data System (PI-RADS) and Likert Scales. Radiology 269, 482-492, doi:10.1148/radiol.13122233 (2013).

Missing Details on Statistics: None that I can see. No issues here.

Missing Citations: The bibliography is appropriate to this study. I do not have any suggestions here.

Additional suggestions: None

Stylistic Issues: None, this is a well written manuscript.
Thank you for that statement.

Reviewer #3:

The team of Prof. Grimm has extended on their groundbreaking work in Cerenkov imaging and present a unique clinical proof-of-concept study with by far the highest number of patients included so far report on a novel imaging system and show images for a range of isotopes that are frequently used in the clinic. By reporting on this data the manuscript substantially advances the scientific and translational potential of Cerenkov imaging.

Major comments:
1. While I find the reported proof-of-concept data a big step forwards for the implementation of Cerenkov imaging, I have my doubts on the valorization aspects presented. I would suggest that the authors tune down their “LightPoint Medical business plan (is the technology a new product in the company’s portfolio?)” in the introduction and discussion and rather focus on the science behind the unique data set that has been collected. This will in my view improve the manuscript further. Note: Before any claims can be made on cost benefit or replacement of PET by CLI imaging, more is needed than mere examples. For example, how do these images impact clinical decision making? I’m sure these aspects are on the radar of this expert team, but I strongly suggest they wait for this data before making claims in this direction.

We thank the reviewer for these comments, wholeheartedly agree and apologize for suggesting that this work was part of a business plan. We have removed the mention of the company in the body of the manuscript and added an emphasis on improving radiology access to patients. The camera used here was constructed entirely with funding from a NIH grant and is not staged for commercial production. We provided a cost comparison, as it is a barrier in global radiology capacity with the intent that a less expensive imaging method could be used where the initial or additional PET/CT scanner may be prohibitive. We have expanded the CLI data in this revision presented to provide numerous examples for each radiotracer. While we suggest how clinical decisions could be changed with this camera, a clinical example still is not planned, and in addition hospitals will adopt this technology at different rates as we state PET/CT still provides the gold standard image. This trial was

designed to show the feasibility of CLI in patients and its possible implications. Further trials will have to address the effect CLI has on decision-making and, as a much higher hanging fruit, possible effect on outcomes. This point has been added to the discussion.

2. Since the authors specifically claim novelty through presenting a novel imaging system, they should extend on their description of the engineering aspects of the system shown in figure 2. Despite the many valuable details mentioned in the supporting information, it is hard to assess the imaging system itself. In endoscopy fiberscopes are considered to suffer from reduced sensitivity. Perhaps the authors can explain in the manuscript why, apparently, this limitation is not an issue for Cerenkov imaging?

We thank the reviewer for these comments and agree, we were positively surprised at the beginning of this trial that we could indeed see clinical Cerenkov. We have expanded on the engineering description of the system and now acknowledge in the manuscript that the addition of the fiberscope has likely limited our sensitivity and provided citations for such. As the CLI system is constructed, removing the fiberscope is not possible without significant disassembly. We assume that the addition of the fiberscope reduces the sensitivity of the device and provided other endoscope citations mentioning this loss on page 14. Even with this likely handicap on CLI sensitivity we can still image all of the clinical radioisotopes. This is likely due to the extended imaging times (up to 15 mins) and the increased sensitivity provided by on sensor binning in our setup to ensure the capture of CL. In expanding the engineering aspects of the system, we have also included two new supplemental figures covering the calibrated sensitivity of the camera down to 0.03pW under both 4x4 and 8x8 binning modes (supplementary figure 4), in addition to a resolution test pattern to show resolution changes with binning (supplementary figure 5). We also added a new tissue phantom with tumors at several depths to correlate CLI observed with depth in Supplementary Figure 11 and found the 600nm filter to provide the best images for the installed filters. As such we have removed the claim about depth determination by filtering but emphasized that the intensity seen in the 600nm filter does correctly represent the activity and depth of lesions seen, though a PET/CT was required to understand the CL intensity.

3. The authors discuss filtering of the emission wavelength and how this can impact clinical utility (Figure 3). But in current form this data seems to be suggestive more than anything. One could potentially get similar images when the sensitivity of detection is reduced independent of the wavelength. Is there a way to exclude this potential artefact and truly validate the effect is wavelength dependent? E.g. by using a rainbow visualization of the signal intensity? Please note: For fluorescence imaging evidence there is increasing emerging that fluorescent emission in the 500-600 range also provides clinical value e.g. fluorescein, suggesting that perhaps the need to shift to high wavelength regions is not perse a requirement for clinical utility.

We thank the reviewer for these concerns. We agree more rigor was needed to define the application of filters. We added in Supplementary Figure 11, a tissue phantom containing identical tumors at different depths and used the filter set to attempt to identify lesion depth. We found that the filters below 600nm were not helpful in determining depth, while the 600nm long pass filter could quantitatively show for two similarly avid lesion/phantom spheres which one was deeper where a PET/CT supported that claim for the 600nm filter. We have also added more citations on page 10 to highlight the clinical use of shorter wavelength probes such as FITC.

4. It seems to me that Figure 5 provides data that could convert into new scientific insights and data

of which the interpretation is instrumental for the future utilization of Cerenkov imaging. I would urge the authors to try and get more information from this data and make stronger conclusions in the results and discussion section based on this information. For example, is there potential clinical value in using the measured signal intensities (and energy of the radiation that is the source of the Cerenkov light) to determine the depth of origin? Or is it very much like clinical fluorescence imaging: you either see something, or you don't? For the latter the lack of quantitative data (and tissue attenuation of signal) limits the clinical utility and essentially require it to be used in combination with a radiological modality. With many of the radiotracers used you can overcome this limitation through e.g. PET imaging or beta-probe based signal detection, please discuss this in more detail.

We thank the reviewer for mentioning this and have looked more closely at the data. Clinical benefits from this intensity analysis include for each radiotracer tested, any region with an intensity two-fold above the noise should elicit a closer look at the image. Namely, does the area imaged likely deviate from physiologic distribution or uptake of the tracer, suggesting the presence of pathologic tissue. In this example, FDG could be used easily to identify superficial lesions in the neck and extremities, providing the patient remains still to minimize physiologic uptake. In the extreme example of thyroid patients receiving Iodine-131 for thyroid ablation therapy, any intensity above the camera noise in the thyroid bed, excluding the salivary glands, would suggest remaining thyroid tissue and warrants further evaluation, and possibly an additional round of iodine therapy. For 177Lu-DOTATATE patients the significant intensity trend would allow physicians to estimate the disease burden in the patients, though it should not be treated as a quantitative metric given the high chance of light attenuation in deeper parts of the body. In short there are first pass clinical decisions that can be made with this information, and should be used to direct patient treatment if no other diagnostic is readily available, depending also on the nature of disease (thyroid and neck adenopathy being easier to assess then splanchnic parenchymal lesions)

Minor comments:

1. Minor: The example images of patients that have been provided in Figures 2 and 4 are impressive. Would it be possible to include more images in the SI? Just so that a reader can better assess how image quality may vary between patients.

We thank the reviewer for the observation, and we have added supplemental figures 8, 9 and 12 to cover more patient images for each isotope, imaging at different angles to understand any change in image fidelity based on projection and additional patient filter examples.

2. Minor: The authors should briefly mention the radiation exposure concept behind Cerenkov imaging. I agree that this is irrelevant when you image patients that were scheduled for PET, but this feature becomes relevant when you propose to use Cerenkov imaging as independent modality for diagnostics or image guided surgery. Mainly because the alternatives would yield less radiation exposure for patients.

We thank the reviewer for this observation and have added a few sentences on page 16 as part of the discussion mentioning that expanding the use of CLI would expose more people to radiotracers, and that the use should be directed to oncology and time sensitive studies.

Rebuttal 2

Point by point response to the reviewers.

Reviewer #1:

The authors addressed prior concerns. Overall, this is an important work given the novelty and proof that this imaging is possible in patients and informative of disease activity. Further work will be needed for clinical use, but this is an important step in the development.

We thank the reviewer for their insightful comments and the help in generating a more meaningful manuscript for readers. We are also glad to see that the importance of this publication is acknowledged and agree that this warrants future work.

Reviewer #2:

Summary: The authors have put a good deal of effort into addressing my concerns of the original manuscript. The revised manuscript is much improved. The new version of this work now has a more pragmatic tone in regards to the utility of this method. Additionally, efforts to better quantify the sensitivity of this new method are marked improvements to this work. The authors have addressed most of my concerns, but I still have 2 major issues at this point that I feel should be addressed.

Degree of Advance and Implications: This revised manuscript does a much better job at justifying the novelty and utility of Cherenkov imaging in regions where PET/CT scanners or SPECT scanners are scarce. The authors bring up a very good point that this would be a useful tool in triage scenarios when deciding if a PET scan is needed. You could also argue that this would be a reasonable approach in any hospital system, even one in a wealthy country, and could help reduce costs to patients.

Thank you for the comment that this could be supplemented in any hospital system. We agree and have added to sentences on line 68, 336-339, and 371 to suggest the broader application of the imaging modality.

Major concerns:

1) In regards to orientation dependence, the authors supply examples in Supplementary Figure 9 of Cherenkov signal vs orientation. I would say that based on these images that there is a significant dependence on camera orientation relative to the patient. For instance, in Example 2, we can see in one case a single focal point for the Cherenkov signal while in the next view there appears to be 2 focal points. SPECT here would indicate only one focal point of uptake. Perhaps recommending a preferred camera orientation would be a good route to go here?

We thank the reviewer for this insight. To address the effect of camera orientation on the Cerenkov signal, we obtained sequential angular dependent images of the 1pW calibrated light source used in Supplementary Figure 4. Imaging was carried out at angles between -

45 and +45 degrees in 15 degree increments. As is common with many optical systems, we found an increase in both intensity deviation, as well as a loss in total signal with increasing angle deviation from planar (0 degree) orientations. We have added this information as supplementary Figure 11 starting on line 239 of the supplement and described the results between lines 233 and 237 of the main text. Additionally, we have amended the manuscript on line 363 to specify directly anterior or posterior for imaging positions, such that scans would anatomically most closely match that seen for SPECT and PET coronal images. We have also specified on lines 236 and 237 that the images were acquired as close to the planar view as possible to maintain image fidelity and intensity uniformity, as permitted by patient comfort.

2) The authors state the "limitations in Xofigo have been pointed out more clearly to address this point" in the current draft, but I do not see where that is. Can you please point to specific language that addressed this? One example is in Supplementary Figure 8E which demonstrates that the perceived uptake when comparing CLI to PET is very different. I do not see these discrepancies discussed in the manuscript.

We apologize for not stating clearly where the additions were made. Limitations of Xofigo imaging can be found on lines 212-215, 219-221, and 782-783, of the main text covering physiologic vs pathologic uptake, and distribution duration. We have also described in the legend of Supplementary Figure 8 on lines 125-126, 142, and 145-146 of the supplement, any possible explanation for gross image discrepancy between the PET and CLI image. These details can be found along lines 162-164, 187-188, 193-195, 219-221 of the main text.

Minor concerns:

1) Looking at Figure 5, it looks like subfigure C has a significant correlation. The P value is lower than in subfigure B. If this is true, can the authors amend their text to reflect this?

We thank the reviewer for the observation and apologize for our omission of significance in Figure 5C. We have amended the figure 5 caption on line 859 to state “Spearman correlations in A, B, C, and G are significant while others did not achieve significance for the patients imaged”.

2) I was able to track down most of the changes, but there were some changes that did not correspond to the locations indicated by the authors in the Response to Reviewers document. For example, Supplementary Figure 5A-B was emphasized for the question on scintillation, but I think the correct figure to refer to here is Figure 7A-B. Also, in many places the indicated pages where changes were made were not correct. For the most part, I was able to track down changes with a search tool, but more accurate page designations would be more helpful in the future.

We sincerely apologize for making it difficult to identify where revisions were made. We have added line numbers in this revision to ensure proper location of edits and also show edits in red.

Reviewer #3:

I would like to thank the authors for thoroughly addressing my comments/concerns.

We thank the reviewer for their insight and improving this manuscript with his/her suggestions.