nature biomedical engineering

PEER-REVIEW INFORMATION

Simultaneous quantitative imaging of two PET radiotracers via the detection of positron–electron annihilation and prompt gamma emissions

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

Thu 15 Sep 2022
Decision on Article nBME-22-1654

Dear Prof Grimm,

Thank you again for submitting to *Nature Biomedical Engineering* your manuscript, "Multiplexed PET for true dual simultaneous and quantitative radiotracer imaging", and for your patience in waiting for the complete set of reviewer reports. As noted in previous e-mail correspondence, the manuscript has been seen by three experts, whose reports (including the two reports that I had relayed earlier) you will find at the end of this message.

You will see that the reviewers appreciate the work. However, they express concerns about the degree of support for the claims, and provide useful suggestions for improvement. We hope that with significant further work you can address the criticisms and convince the reviewers of the merits of the study. In particular, we would expect that a revised version of the manuscript provides:

* Extended quantification of the phantom experiments, also using clinically relevant ratios of isotope amounts, as per the relevant comments of all reviewers.

* A more convincing study of the advantages of mPET in quantitatively imaging nanoparticle delivery to tumours, as also alluded to by all reviewers.

* Discussion of the limitations of the implementation of multiplexed PET.

* Thorough methodological reporting.

* Substantial rewriting of the manuscript, and enhanced figure design and presentation, to aid understanding (regardless of whether the background of a future reader is radiological physics, or preclinical or clinical

imaging). Reviewer #2 offers useful suggestions. Please make good use of the Methods section, where formulae can be included.

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.

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 20 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 hope that you will find the referee reports helpful when revising the work. Please do not hesitate to contact me should you have any questions.

Best wishes,

Pep

__
Pep Pàmies
Chief Editor, *Nature Biomedical Engineering*

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

Some major concerns from this reviewer are listed below:

* Multiplex means more than 2, rather than just 2. That was what it refers to in the optical or photoacoustic world. Currently there is no more than 2 in an experiment, hence I don't think multiplex is accurate terminology for a Nature family journal.

* In the end "providing a rainbow of targets much richer than our current state of single colour PET imaging" is overclaiming. The authors should first demonstrate 2 color PET in the same image, and perhaps need to show multi-color PET in one image to claim that.

* In Figure 4, target expression level in the tumor tissue before and after therapy should be provided to

confirm that PET data truly reflects target expression level. This comment applies to all the experiments. The manuscript reads like from a chemist/physicist/engineer, without much attention to the actual biology behind, or the presence of all intended targets in the tumor tissue and other normal tissues. They took too much for granted.

* Quantitative %ID/g data in the tumor should be shown in most of the figures to enhance readability, especially since the authors claim" quantitative radiotracer imaging" in the title. They need to deliver what they claim.

* Figure 5 is not well-designed. A "drug" loaded onto a nanoparticle that quickly gets released is not ideal. How about a drug that was stably loaded and dual-labeled? The scale bar is quite meaningless as min/max and 0-25 with no unit means nothing (that is the case for most figures, unfortunately). The PET images need to be displayed on the same scale bar to have meaningful visual comparison. Again, quantitative data is much preferred over qualitative data, otherwise they can not claim quantitative in the title.

* For all the tracers, nanomaterials, etc. there was no characterization or QC presented. Are we supposed to just take the authors' words for those (purity, specific activity, etc.)? I understand that some were published in previous papers, but this manuscript by itself should be self-contained and complete. For commercially available agents or those that are GMP and in human studies etc, they are not necessary as long as proper papers/sources are cited. But for all those "In house synthesized" ones, some data would be helpful to enhance rigor and reproducibility.

* Can this method be used to show, in one image, the separate distributions of two tracers, like dual-color SPECT or CT? i.e., if two separate images are made with two different color scales, if the images are then overlayed, does the product image provide any utility or is it difficult to interpret? This question can be answered with a demonstration in the manuscript.

* Computation time is a significant trait of any reconstruction algorithm. What is the cost of this reconstruction regarding file size, data management, reconstruction time, etc., when compared to regular PET? Some tabulated summary would be nice to see.

* Dose of each isotope used in the liver image (Figure 3) is apparently about half of what would typically be administered. Of course, this prevents excess dose, but also sacrifices some image quality. It is said that "mPET provided increased image information density with little bias (<10%), similar spatial resolution and contrast, and a small relative noise increase (<10%) compared with that obtained with single tracer PET acquisitions." Can you provide more specific values for these quantities, and perhaps a justification? In principle, these values depend on the isotopes used. Is there a way to theoretically determine the increase in noise when two isotopes of a given half-life and activity are imaged with mPET vs. regular PET?

* In preclinical studies, activity of harvested organs is often measured in a device separate to the PET scanner in order to confirm PET-recorded activities. With two isotopes, what is the method with which one can confirm PET-recorded activities? How do we confirm that both isotopes can be measured accurately (e.g. via dose calibrator and/or PET)?

* If sequential PET entails more dose for the patient, what is preventing this from being tested clinically in an application for which this method of reconstruction is useful, and how can this goal be made a reality?

Some minor concerns are listed below:

* Some incorrect nomenclature was in the text, for example Figure 6, PSMA is the target, not the tracer. Similar mistakes occur many times in the text, figure, legend, throughout.

* Some text such as "Nanoparticle Drug Delivery with mPET" is misleading. mPET does not do drug delivery.

* Figure 1 legend "(blue ß+ emission)" is scientifically inaccurate. The authors should not confuse positron and 511 kev gamma rays.

* abcd and ABCD were both used in figures and legends. Need to be consistent.

* In SI, it was I-131, not I-124. Is that correct?

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

The authors use a triple-based reconstruction to image the biodistribution of positronic radioisotopes emitting one or more prompt gamma rays in addition to the 511 keV gammas of positronic annihilation. Prompt gamma emission concerns mostly metal-based “exotic” radioisotopes, but also radio-halogenates (76-Br, 120- and 124-I). The prompt gamma is registered as an additional event to the two positron annihilation gammas. Hence, the term triple since three lines of response are detected simultaneously: one connecting the 2 annihilation gammas, and 2 connecting the prompt gamma with each of the annihilation gammas.
The paper has a very strong fundamental and methodological basis. The technology has been applied to a significant number of situations using different scanners and 5 phantom and animal studies. It is greatly original with respect to the state of the art. Implications of the work may turn out to be important in the field of molecular imaging.

The idea to use prompt gamma rays to separate signals from two radioisotopes in a single PET session is quite appealing because it offers simultaneous co-registered readings of multiple radiotracers, hence, of multiple targets. One important advantage of the “Multiplexed PET (mPET for short) is, therefore, its capacity to acquire synchronized and diachronic information. For instance, an important point that could be addressed by mPET would be, “is the drug hitting the target lesion?”, a major question in cancer therapy. One may also anticipate that mPET could provide readings of duplex fundamental information in oncology, cardiology, and neurology, using different tracers for various receptors and transporters. All these arguments make the study by Pratt and co-authors interesting.

Naturally, other combinations of imaging techniques have similar ambitions, e.g., SPECT as mentioned by the authors, PET-MRI, Multiplexed MRI, etc. The advantage of mPET is its very high molecular sensitivity presently unmatched by any other technique, and the crucial point is that it is a purely software based application that could be transferred very rapidly to PET since it does not need any developments in Physics and instrumentation. The authors should be commended for this exploration potentially of high interest.

Essential comment
Figure 3 liver phantom experiment: Do I get it right that the lesion to the background (LTB) is circa 200? This is not realistic in clinical practice. Unfortunately, this fogs the confidence in the mPET method. A series of realistic ratios should be given throughout the manuscript (this can be inserted as supplementary data). I strongly recommend the authors provide all the information on 2-tracer ratios, not only in the images but as numbers reporting true coincidences and prompts and ratios of radiotracer concentrations using their iterative method treatment.

Major comment
My major criticism of the manuscript is that its construction hesitates between a demonstration of the physics and math of mPET reconstruction, and a plead for the new biological information that may be obtained with mPET. This is far from ideal, as both the reader somewhat familiar with the physics of PET is frustrated by the lack of details in the presentation of the method, while the biologist is merely given some proofs-of-concept that mPET may address biological questions, but without a real eye-catching data in the results (and with some frustrating unnecessary data, i.e., Fig. 5). Without going as far as to pretend that Physics and Biology can be made crystal clear in the same paper, I strongly recommend providing in a simple apprehensible manner the key points raised by mPET.

In that respect, the key point of the right panel “Iterative image reconstruction” in Figure 2 is poorly informative and should be summarized in the main text. The specific details for isotopes and cameras can be given in Methods (lines 440+) where they should be presented in a structured manner and detailed.

Conversely, biological experiments should be exposed in a logical manner. There is a progression from the phantom (Fig 3) to the therapy monitoring (Fig 4) to nanoparticle delivery (Fig5) to CAR-T cell therapy (Fig6). It is essential here to provide intermediate conclusions in order to inform the reader with the potential advantages and disadvantages for each application.

Special questions and comments

1. Since the emission of the prompt gamma has no directionality, it can reach any detector. In their method, the authors use only two lines of response (LORs), the annihilation coincidence and one of the two prompt-annihilation LOR. Figure 1 and Supp. Fig.1 both illustrate a case in which the second prompt-annihilation LOR is obviously not crossing the subject and will be easily rejected, but what happens if this is not the case and if both prompt-annihilation LORs across the subject? Are these events rejected by the reconstruction algorithm or do they pollute the V-shaped LOR? Calculating the probability that this may happen and calculating of its effect on the quality of the reconstruction, would be appreciated.

2. Information should be made clear about the difference in spatial resolution obtained with and without prompt correction. Is the resolution improved for a single prompt emitting radioisotope?

3. Table 2 shows that the ratio of prompt/positron emissions varies according to the radio-isotope, from 13% for Rb-82 to 100% (Sc-44), and 124-I and 68-Y used here for antibody-labeling having ratios of 51% and 33/83%, respectively. In other words, using 124-I alone would yield two times more LORs than VLORs. How is mPET coping with that? Are the annihilation-only LORs rejected? If not, will they not pollute mPET reconstruction and preclude separation from pure beta+ radio-isotopes?

4. It is not clear what were energy detection ranges used with the different instruments and how it was optimized. Use precise values and not vague terms such as “can be expanded (line 448)”. How were the various photopeaks matched to account for detection sensitivity at different energy windows?

5. How were the accepted/rejected triple coincidences chosen? Supplementary information lines 148+ should show quantitative data.

6. The argument about avoiding double CT irradiation is of limited value since most modern PET-CT use low-irradiation CTs. In my view, the synchronization of the images is a much stronger argument and the authors should rather insist on that point.

7. How will the authors manage the different kinetics variations among patients? Among tumors? Among tumor drug delivery? Isn’t the optimization of the injection-imaging delay an unsolvable nightmare?

8. There is little interest stemming from the nanoparticle drug delivery experiment with mPET. Here the release of the drug from the nanocarrier is completely obvious and could be made evident with other techniques. Isn’t there a better example of using slow drug-release nanoparticles?

9. Figure 4 and supp. Fig.2: How do you explain the higher FDG uptake under BRAF or MEK therapy?

10. Supplementary Fig 3: not the same mice, quantitative values must be provided.

11. Supplementary Figure 4: it’s very hard to read the figures with this color scale. Again, numbers would be necessary for data interpretation.

12. Line 126:” expanded energy range” … please give numbers.

13. References are not mentioned properly, abbreviations are not standard or absent and some titles (#16) have been omitted.

14. Please correct typos and English spelling, e.g., foot pad and not food pad, etc. Color scales of the supplementary data are barely readable. Some panel references (Fig. 6 are not indicated in the figures.

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

This paper innovatively proposed a general multiplexed PET reconstruction method (mPET) to separate the two tracers with single acquisition, relying on the triple coincidences of some specific radioisotope due to the additional prompt gamma rays. To validate the proposed approach, several preclinical/clinical experiments were designed and performed well on phantoms or mice, and the results showed two tracers could be differentiated successfully. The paper is well-organized and clear to read with abundant supportive data and

nice images, however, I still have some questions on the described methods and experiments as follows:

Major:
1. From L84 – L104, you gave a brief intro of the new general reconstruction method, and it seems like mostly inspired by the 25th and 29th refs, and explained more in the supplementary files. Considering the algorithm is kind of important for readers to understand the entire method, maybe it is better to re-organize the methods section (e.g. including more mathematical expressions) and present it before the results.

2. For the preclinical phantom experiments, a 350-700 keV energy window was chosen during the acquisition, how was this parameter optimized? Especially, the emitted prompt γ-rays have different energies? Does it require to redesign the PET detectors to detect high-energy gamma rays more efficiently?

3. In Figure 3, the 2th cavity in C seems kind of weaker, especially at the edges, why caused this? Is it because the recon or other reasons?

4. In Figure 3, The E is standard PET, if the 124I signal could be attenuated as scatter events, then why is the 89Zr region so dark in E and bright in F? Do the E, F and G share the same colorbar?

5. For the second experiment, the [124I]I-trametinib injected 24 hours before the [18F]FDG, considering the ~4 days half-life of 124I, it is reasonable and the results also validated the sampling protocol. Have you ever tried different injection time of the second tracer, or simultaneously? Do they also work well with mPET method?

6. For the nanoparticle drug delivery, do you think the standard PET recon could approximately seem to be the sum of [124I]I-trametinib and [89Zr]Zr-ferumoxytol? In Figure 5, images in the first row seem stronger than the sum of corresponding mPET images in the second and third rows. Anyway, it is better to rescale and make sure the images are comparable, so we could explore more about it.

7. From biological effect point of view, the emitted high-energy γ-rays (more than 511keV) is not desired. On the other hand, PET detectors are designed to detect the pure 511 keV emitter. Therefore, there are some efforts for separating dual-tracer signals from reconstruction point of view. The authors have to address the limitations of the proposed method.

Minor:
1. In Figure 3, A-G is better matched with the annotations a)~f).
2. In Figure 4, A-D is better matched with the annotations a)~d).
3. In Figure 5, the front size of the mice images looks too small, please check it.
4. In figure 6, I just noticed the colorbars of 68Ga-PSMA and 124I images have black outlines, maybe the unseparated series should keep the same style.

Wed 01 Feb 2023
Decision on Article nBME-22-1654A

Dear Prof Grimm,

Thank you for your revised manuscript, "Multiplexed PET for true dual simultaneous and quantitative radiotracer imaging". Having consulted with the original reviewers (whose 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*.

We will be performing detailed checks on your manuscript, and in due course will send you a checklist detailing our editorial and formatting requirements. You will need to follow these instructions before you upload the final manuscript files. In the meantime, please consider the minor points from Reviewers #1 and #2.

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

Best wishes,

Pep

__
Pep Pàmies
Chief Editor, *Nature Biomedical Engineering*

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

I have read the manuscript and the authors' rebuttals. The reviewers' comments were definitely necessary and mostly well-answered by the authors.

I think this paper is getting quite robust, with a few pending issues with:

1. I still think “Duplexed PET” is more accurate than mPET for this paper. Given what was demonstrated here, it is scientifically more accurate to name this one dPET and the next paper (where they can demonstrate multiple colors or rainbow of colors with better PET scanners) mPET. One needs to learn to walk before they can run.

2. As suggested by reviewer 1, point 5, the authors should look at a case study where the drug-particle combination does work. They didn't do this, nor did they say why they wouldn't do this. Although, I accept the authors' justification for having (and keeping in the final draft) a case study where the combination doesn't work.

3. The authors' treatment of Figure 5 and Supplementary Figure 8 is still not great. Max and min % are still not stated and a comparison between the standard PET and mPET (should be dPET) with different scale bars is not straightforward as the figure may imply at first glance.

The responses to much of the comments were discussion/arguments, albeit politely. Other authors may have done significantly more experiments or made more concrete changes to the manuscript.

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

My opinion is that the authors addressed all of my comments and made the changes and additions accordingly.

Minor changes to improve the manuscript:

line 158, a word is missing

line 123 please define Wv1 and Wv2 in the text.

line 505 It would be nice to comment on Table 2 by indicating which prompt-emitting isotopes could be used for mPET on any PET scanner without modifying the instrument, and which would require improved energy resolution, etc. The reader interested in the new technology could take a glimpse at the kind of double tracer experiments he could plan with respect to the domain of imaging (cancer, brain, cardiovascular...

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

In this revised version, the authors have addressed my concerns.

Rebuttal 1

REVIEWER #1:

We want to thank the reviewer for their thorough review. The major concerns from this reviewer are addressed below:

1. Multiplex means more than 2, rather than just 2. That was what it refers to in the optical or photoacoustic world. Currently there is no more than 2 in an experiment, hence I don't think multiplex is accurate terminology for a Nature family journal.

We thank the reviewer for their comment and the point is well taken. As the algorithm has been described elsewhere and is capable of more than two (the PET scanners we have do not have good enough energy resolution/information and therefore they are currently limited to two isotopes currently), we would like to keep the multiplex PET name. Broadly the term multiplex refers to the processing of layered information from a single data-stream. As the two isotopes are extracted from the single listmode, we feel the multiplexed PET term is appropriate. While we acknowledge where the reviewer is coming from, we would prefer to keep the and the name.

2. In the end "providing a rainbow of targets much richer than our current state of single colour PET imaging" is overclaiming. The authors should first demonstrate 2 color PET in the same image, and perhaps need to show multi-color PET in one image to claim that.

We have toned down the phrase, though still kept the general outlook in the conclusion that additional colors could be done. To provide more than a two-color PET image, new PET systems with good energy resolution and adapted data acquisition systems would be required. Those additions would detract from the message this manuscript is trying to show; that data from current preclinical and clinical PET systems can utilize a new PET reconstruction method (without any hardware or acquisition software modification) when a triple emitting isotope is paired with a double emitting isotope.

3. In Figure 4, target expression level in the tumor tissue before and after therapy should be provided to confirm that PET data truly reflects target expression level. This comment applies to all the experiments. The manuscript reads like from a chemist/physicist/engineer, without much attention to the actual biology behind, or the presence of all intended targets in the tumor tissue and other normal tissues. They took too much for granted.

We thank the reviewer for the comment to show and confirm expression level. We chose many of the radiotracers in the main figures as they have been previously published and validated as targets. We understand that the trametinib imaging in figure 4 is relatively new, and IHC could be provided if requested, though the target is endogenously expressed. The other radiotracers used in the main figures, however, have been extensively studied, and we feel expression confirmation is beyond the scope of this paper. The goal of the manuscript is to show biologist, engineers and chemists alike how to increase PET information density, while simultaneously providing known biology examples to others to highlight the possibility of dual isotope imaging for their new studies. We have added this statement in the introduction on lines 90-91 to reflect this message.

4. Quantitative %ID/g data in the tumor should be shown in most of the figures to enhance readability, especially since the authors claim" quantitative radiotracer imaging" in the title. They need to deliver what they claim.

We thank the reviewer for this comment, and we sincerely apologize that our figures were not up to the quality they should have been. All images showing the individual tracers as separated are calibrated and presented as injected dose per cubic centimeter. We have left the unseparated images uncalibrated, first as the standard reconstruction does not properly include all the triple events as they cannot be reconstructed, hence would inaccurately represent the quantitative power of PET. Second, each isotope has different decay and positron efficiencies, making a single calibration factor difficult to implement for both isotopes over longitudinal studies.

5. Figure 5 is not well-designed. A "drug" loaded onto a nanoparticle that quickly gets released is not ideal. How about a drug that was stably loaded and dual-labeled? The scale bar is quite meaningless as min/max and 0-25 with no unit means nothing (that is the case for most figures, unfortunately). The PET images need to be displayed on the same scale bar to have meaningful visual comparison. Again, quantitative data is much preferred over qualitative data, otherwise they cannot claim quantitative in the title.

We thank the reviewer for the comment and agree that a loading success story looks much more appealing in a manuscript figure. However, we want to point out that in this case it was exactly this study that demonstrated to us that the drug particle combination did not work well and provided an important example how multiplexed PET provided valuable information even if it is not a success story. Further, finding radiolabeled drugs that can be synthesized with a bromine or iodine halogen for mPET, while not modifying the drug itself is not trivial. The purpose of this experiment was to show that mPET could easily identify the fate of the two drugs, and that this can be quite different from what in vitro loading stability studies suggest.

We again apologize that the scale bar was not properly annotated as the separated images provided already are calibrated to injected dose per gram or cubic centimeter. We have updated all figures with calibrated separated images. See also Reviewer 2 Major Question for a similar discussion on Figure 5 merits.

6. For all the tracers, nanomaterials, etc. there was no characterization or QC presented. Are we supposed to just take the authors' words for those (purity, specific activity, etc.)? I understand that some were published in previous papers, but this manuscript by itself should be self-contained and complete. For commercially available agents or those that are GMP and in human studies etc., they are not necessary as long as proper papers/sources are cited. But for all those "In house synthesized" ones, some data would be helpful to enhance rigor and reproducibility.

We apologize for this oversight on radiotracer characterization and have extended this information now significantly. Since the PET imaging was at the core of the manuscript this has been neglected. We have now assembled Supplementary Fig. 17, covering instant thin chromatography (ITLC) of the antibodies and nanoparticles, and HPLC radiotracers for the small molecule [^{124}I]I-trametinib and peptide [^{86}Y]Y-DOTA-PSMA. We have included

radiochemical purity, specific activities, improved the relevant methods for the characterization, and separation with examples of free isotope movement for the ITLC samples. We did not include characterization details for [^{68}Ga]Ga-PSMA-11 or [^{18}]FDG examples as these materials were readily available in kit form and/or provided from the clinical radiopharmacy and described in both the methods section and Tables 1 and 2.

7. Can this method be used to show, in one image, the separate distributions of two tracers, like dual-color SPECT or CT? i.e., if two separate images are made with two different color scales, if the images are then overlayed, does the product image provide any utility or is it difficult to interpret? This question can be answered with a demonstration in the manuscript.

We thank the reviewer for this interesting question. Upon this question, we actually explored some methods to improve the visualization of the results, and to make better use of the multi-channel information collected. One approach, related to the suggested product image is the computation of the geometric mean of the two images. This is known to produce a result that enhances the differences between the two images respect to the most common arithmetic mean [PMC3127352]. We have included videos of overlapped [^{68}Ga]Ga-PSMA-11 and ^{124}I images (using two color scales), as well as Supplemental Figure 12 showing the geometric mean (i.e. square root of their product).

8. Computation time is a significant trait of any reconstruction algorithm. What is the cost of this reconstruction regarding file size, data management, reconstruction time, etc., when compared to regular PET? Some tabulated summary would be nice to see.

We thank the reviewer for this indeed important question. The computation time of our implementation has not been optimized yet, and we were not looking to produce a faster reconstruction method, but rather one to allow for two isotopes. The list-mode data processing currently takes about 5 minutes for a 2Gb file in 1 CPU. Then we reconstruct the generated double and triple sinograms using an adapted version of our image reconstruction code written in C++/CUDA called GPU-FIRST. We used 60 iterations with no subsets, and it requires 5 minutes per sinogram (i.e. 10 minutes in total for double and triple sinograms). In this implementation, we perform an initial reconstruction with only attenuation and randoms correction (i.e. without scatter and background correction), and then use the simulator MCGPU-PET for the estimation of scatter and background. Therefore, the total reconstruction time is about 30 minutes per acquisition. Of course, this may be significantly reduced by using subsets, estimating the scatter and background in the middle of the reconstruction (as it is commonly done in clinical scanners), instead of performing two reconstructions and performing the list-mode data processing with multiple CPUs. These optimizations are planned for future work as they will be helpful for other users. We have added these details to the introduction starting at line 93 and added the above section to the methods section on line 548.

9. Dose of each isotope used in the liver image (Figure 3) is apparently about half of what would typically be administered. Of course, this prevents excess dose, but also sacrifices some image quality. It is said that "mPET provided increased image information density with little bias (<10%), similar spatial resolution and contrast, and a small relative noise increase (<10%) compared with that obtained with single tracer PET acquisitions." Can you provide more specific values for these quantities, and perhaps a justification? In principle, these values

depend on the isotopes used. Is there a way to theoretically determine the increase in noise when two isotopes of a given half-life and activity are imaged with mPET vs. regular PET?

This is a great observation. Indeed, determination of the impact of simultaneous dual acquisitions with mPET on the image quality is not straightforward. Since we use iterative reconstructions and guided filters, the relation between the signal-to-noise ratio in the data (with either mPET or regular PET) and the noise in the final images, is not linear. Using a naïve approach with analytic reconstruction (FBP) with a simple subtraction and not including the guided filter and the separation of images applied for mPET, it is possible to estimate the impact on the noise level in mPET vs regular PET. However, these analytic results will be significantly worse than the ones obtained with our method. We have therefore included Supplementary Fig. 4 with a study on this where we reduced the reconstructed counts and observed no deviation until below 20% of the triple events. This deviation was due to the number of triples approaching that of the triple noise. Furthermore, we are currently working on further reducing increased noise with Deep Learning methods which have shown to be very effective in reducing the noise in PET image reconstruction. With respect to the <10% bias and noise increase values, they were obtained from phantom measurements (such as the ones shown in Fig. 3 for the Inveon and mCT) and Monte-Carlo simulations.

We conducted additional experiments testing the performance of mPET, examining measured noise for each reconstructed, image compared with the standard or mPET method, as well as a linearity test using only a triple emitting isotope. These experiments have been added as Supplementary Figs. 2 and 3. In the noise analysis as part of Supplementary Fig. 2d, by ROI analysis, the standard and mPET methods had comparable noise values for the liver which we mention in the figure legend, where 20.7 % was measured for the standard reconstruction, while 17.9 % and 18.0 % noise were observed for the doubles and triples as reconstructed with mPET. The noise observed for the triples without the mPET process would be 45.7 %. In the resolution analysis, double and triple events could be reconstructed without loss of resolution compared to the standard reconstruction method. Furthermore, we have included the probability of doubles, triples (with 2 and 3 LORs within the field of view) for reference. Third, we did a triple event linearity analysis, showing that the abundance of triple events was linear over the entire range of the scanner, with less than a 5% deviation calculated for all of the activity amounts used. Most important was that any crosstalk observed between wells in the phantoms was also observed with the standard reconstruction, showing the similarity of the method results despite being different methods. Overall, we feel these additional data plus the improved phantoms in Figure 3 greatly improve the quantitative description of the mPET method. See also Reviewer 2 Question 2 for a similar question of spatial resolution.

10. In preclinical studies, activity of harvested organs is often measured in a device separate to the PET scanner in order to confirm PET-recorded activities. With two isotopes, what is the method with which one can confirm PET-recorded activities? How do we confirm that both isotopes can be measured accurately (e.g. via dose calibrator and/or PET)?

We thank the reviewer for this question. Indeed, terminal biodistribution is considered an orthogonal approach to confirm activity present in the tissue measured by gamma counting, provided a ratio can be established to accurately separate both isotopes. A dose

calibrator could be used for measuring the activity of multiple isotopes in a sample based on the differences in their half-lives. This simply requires several measurements at different time points and use the differences in their decay to obtain the separated activity of each isotope. Furthermore, PET images can be calibrated using phantom acquisitions, where the known activity concentration of each isotope is known. Figure 3d shows the critical experiment by revealing that when known amounts of activity from a dose calibrator are added to the PET and imaged, mPET accurately recovered the activity present.

11. If sequential PET entails more dose for the patient, what is preventing this from being tested clinically in an application for which this method of reconstruction is useful, and how can this goal be made a reality?

We sincerely thank the reviewer for this question and are planning for such a study. In this paper, we have already shown that a clinical scanner is capable of mPET imaging. However, we would like to point out that a clinical study is not a trivial enterprise. It requires existing funding to even submit a study administratively and then additional funds to execute the trial (for scanner time, research assistants, nurses and tracers) in addition to significant time to move the study through the various research committees and then the institutions' IRB for the administration of two radiotracers simultaneously. IRB approval and funding would take significant time beyond the scope of this paper.

Some minor concerns are listed below:

* Some incorrect nomenclature was in the text, for example Figure 6, PSMA is the target, not the tracer. Similar mistakes occur many times in the text, figure, legend, throughout.

We have corrected the manuscript to ensure clear description of the PSMA target and the molecular imaging agent being used to target PSMA.

* Some text such as "Nanoparticle Drug Delivery with mPET" is misleading. mPET does not do drug delivery.

We thank the reviewer and have amended the section title to reflect "Imaging Nanoparticle Drug Delivery with mPET"

* Figure 1 legend "(blue ß+ emission)" is scientifically inaccurate. The authors should not confuse positron and 511 kev gamma rays.

We apologize for the confusion and have revised the figure legend on lines 357-368 to explain better PET imaging and the origin of both positron, 511 annihilation photons and prompt gamma.

* abcd and ABCD were both used in figures and legends. Need to be consistent.

We apologize for the inconsistency and have updated figures and legends to a single lowercase setting.

* In SI, it was I-131, not I-124. Is that correct?

The reviewer is correct the loading experiment was done with ^{131}I-trametinib while all imaging studies of course were conducted with ^{124}I-trametinib. While different isotopes, the trametinib isotope variants are expected to perform as the stable ^{127}I-trametinib since it is the same chemical element.

REVIEWER #2

The authors use a triple-based reconstruction to image the biodistribution of positronic radioisotopes emitting one or more prompt gamma rays in addition to the 511 keV gammas of positronic annihilation. Prompt gamma emission concerns mostly metal-based “exotic” radioisotopes, but also radio-halogenates (76-Br, 120- and 124-I). The prompt gamma is registered as an additional event to the two positron annihilation gammas. Hence, the term triple since three lines of response are detected simultaneously: one connecting the 2 annihilation gammas, and 2 connecting the prompt gamma with each of the annihilation gammas. The paper has a very strong fundamental and methodological basis. The technology has been applied to a significant number of situations using different scanners and 5 phantom and animal studies. It is greatly original with respect to the state of the art. Implications of the work may turn out to be important in the field of molecular imaging.

The idea to use prompt gamma rays to separate signals from two radioisotopes in a single PET session is quite appealing because it offers simultaneous co-registered readings of multiple radiotracers, hence, of multiple targets. One important advantage of the “Multiplexed PET (mPET for short) is, therefore, its capacity to acquire synchronized and diachronic information. For instance, an important point that could be addressed by mPET would be, “is the drug hitting the target lesion?”, a major question in cancer therapy. One may also anticipate that mPET could provide readings of duplex fundamental information in oncology, cardiology, and neurology, using different tracers for various receptors and transporters. All these arguments make the study by Pratt and co-authors interesting.

Naturally, other combinations of imaging techniques have similar ambitions, e.g., SPECT as mentioned by the authors, PET-MRI, Multiplexed MRI, etc. The advantage of mPET is its very high molecular sensitivity presently unmatched by any other technique, and the crucial point is that it is a purely software-based application that could be transferred very rapidly to PET since it does not need any developments in Physics and instrumentation. The authors should be commended for this exploration potentially of high interest.

Essential comment:
Figure 3 liver phantom experiment: Do I get it right that the lesion to the background (LTB) is circa 200? This is not realistic in clinical practice. Unfortunately, this fogs the confidence in the mPET method. A series of realistic ratios should be given throughout the manuscript (this can be inserted as supplementary data). I strongly recommend the authors provide all

the information on 2-tracer ratios, not only in the images but as numbers reporting true coincidences and prompts and ratios of radiotracer concentrations using their iterative method treatment.

We thank the reviewer for this critical insight. In the clinical phantom the activity concentration ratio of lesion to liver of ~200 is correct. It was not the intention in the phantom to provide a clinical example or such a high ratio, but rather that the same activity in the field of view could be separated on a clinical scanner. The high ratio was not necessary for the mPET to work. The reconstruction method relies on differentiation between triple coincidences and the traditional doubles, where image reconstruction is best. Indeed, the preclinical phantom as well as the other preclinical examples use comparable activity amounts between radiotracers as well as amounts used in previous works and commonly applied. We have added relevant information for each radiotracer production and in each figure legend the amount injected. Information on general two tracer ratios is provided in the main article in Tables 1 and 2, and our noise analysis in Supplementary Fig. 2d includes the noise for the standard reconstruction and mPET (lines 49-52 in Supplement). The sensitivity analysis in Supplementary Fig. 4 shows the method is stable over a wide range of coincidences and prompts, with only a deviation when the triples approach that of random triples. We have included in the methods section for the clinical mCT the total counts and scan duration as requested on lines 610-612.

Major comment:
My major criticism of the manuscript is that its construction hesitates between a demonstration of the physics and math of mPET reconstruction, and a plead for the new biological information that may be obtained with mPET. This is far from ideal, as both the reader somewhat familiar with the physics of PET is frustrated by the lack of details in the presentation of the method, while the biologist is merely given some proofs-of-concept that mPET may address biological questions, but without a real eye-catching data in the results (and with some frustrating unnecessary data, i.e., Fig. 5). Without going as far as to pretend that Physics and Biology can be made crystal clear in the same paper, I strongly recommend providing in a simple apprehensible manner the key points raised by mPET.

We sympathize with the frustration of bridging between physicians, engineers, chemists and biologists. Our intention was to provide a first step and a broader description of an approach that could interest most reviewers and to exemplify this with a selection of a few relevant biological questions and example cases. As to Figure 5, while it may seem frustrating and a negative result, it nevertheless demonstrates clearly that mPET can provide important information. We did not expect this distribution in vivo based on our in vitro stability studies and subsequently changed our approach with this drug going forward (not part of this paper). We therefore felt that it is important to include this study as it is probably more important than just to highlight confirming studies as science doesn't always produce the anticipated results. Figure 5 is not a success story but quantifying both nanoparticle and loaded drug is extremely important for industry and academia alike. Too often nanoparticle delivery success is based on imaging only one of the two components. We would like to provide this as an example of a true control in the nanoparticle realm, where mPET can provide clear quantitation of nanoparticle delivery (if successful).

However, we have in this revision added the numerical analysis into the introduction before the results to give a better mathematical basis to the work. We want biologists to utilize the method to improve their studies, while we also want mathematical improvements to the method to increase multiplexing, reduce isotope dose, reduce the noise level, etc.. More to the point, multiplexing was not designed to be eye catching and flashy, but rather improve the efficiency of a system that is monochromatic and frankly binary in analysis for positive or negative uptake with a known radiotracer. See also Reviewer 1 Question 5 regarding Figure 5.

In that respect, the key point of the right panel “Iterative image reconstruction” in Figure 2 is poorly informative and should be summarized in the main text. The specific details for isotopes and cameras can be given in Methods (lines 440+) where they should be presented in a structured manner and detailed.

We apologize as the previous comment make bare, we are trying to bridge interests of physicists and biologists alike. We have included in the introduction a more in-depth mathematical representation of the algorithm and have added several supplemental figures (SI 1-4) to better characterize the mPET reconstruction performance in linearity, resolution, noise as well as a more complete workflow. With our data availability statement, we can provide more in depth the code used as needed. We have improved the methods section to include more quantitative information for each radiotracer used and supplementary tables provide specifics on the reagents used. Our intention was not to compare specifics of particular PET scanners, since the method does not rely on one in particular, but that reconstructable events (and triple coincidences) are pulled from the listmode as single or consecutive double coincidences for multiplexing, which is possible with almost all commercially available scanners, preclinically as well as clinically.

Conversely, biological experiments should be exposed in a logical manner. There is a progression from the phantom (Fig 3) to the therapy monitoring (Fig 4) to nanoparticle delivery (Fig 5) to CAR-T cell therapy (Fig 6). It is essential here to provide intermediate conclusions in order to inform the reader with the potential advantages and disadvantages for each application.

We thank the reviewer for this observation and have included a few lines in each section for a brief conclusion of the findings and outlooks.

Special questions and comments

1. Since the emission of the prompt gamma has no directionality, it can reach any detector. In their method, the authors use only two lines of response (LORs), the annihilation coincidence and one of the two prompt-annihilation LOR. Figure 1 and Supp. Fig.1 both illustrates a case in which the second prompt-annihilation LOR is obviously not crossing the subject and will be easily rejected, but what happens if this is not the case and if both prompt-annihilation LORs across the subject? Are these events rejected by the reconstruction algorithm or do they pollute the V-shaped LOR? Calculating the probability that this may happen and calculating its effect on the quality of the reconstruction, would be appreciated.

This is an interesting point. The number of triples with the 3 LORs within the field-of-view (FOV) is typically below 5% with respect to the total triples. For instance, in the case of the Inveon PET/CT scanner (Siemens) used here with its FOV of 100 mm and a bore diameter of 161 mm this amount is about 2.5% of the triple incidents (triples). In this work, triples with 3 valid LORs are rejected. As the 3 LORs would be located at the edges of the FOV (where no activity is expected), neglecting their contribution has a small impact. The extension of the proposed method to include events having the 3 possible LORs within the FOV is possible, but it would not significantly increase the sensitivity while at the same time making the analysis more complex. We have included Supplementary Fig. 2 with the probability distribution in a transverse plane (centered in the axial FOV) of detecting positron-gamma emissions with 1, 2 and 3 detected LORs within the FOV.

Supplementary Fig. 2 shows the sensitivity analysis in a transverse plane centered in the axial FOV of the Inveon scanner for a positron-gamma emitter. The diameter of the FOV in this scanner is 100 mm. The sensitivity distribution of triples with 2 LORs is similar to the double coincidences, although the ratio of sensitivity is about 4% in this case. The triples with 3 valid LORs appear only at the edges of the FOV and their relative importance is quite limited.

2. Information should be made clear about the difference in spatial resolution obtained with and without prompt correction. Is the resolution improved for a single prompt emitting radioisotope?

The background in the acquired data is considered to be the combination of Randoms, scatter and prompt-gammas contributions. Prompt-gamma correction is obtained in a very similar way to the scatter correction: First, an initial image is reconstructed without scatter and prompt-gamma corrections. That initial image is used as the activity distribution to estimate the scatter and prompt-gamma contributions, and a second reconstruction is performed with these corrections. This step may be iterated further (using the last reconstructed image to obtain a more accurate scatter and background correction). In this work, we used 2 iterations of this procedure to make sure that the scatter and background were properly estimated. We observed that more iterations were not needed (i.e. the last background estimation did not differ significantly from the previous one). This procedure is similar to how scatter correction is performed in commercial scanners.
We have added Supplementary Fig. 2 using a ^{124}I acquisition with and without prompt-gamma correction (both with scatter correction) to show how prompt-gamma correction improves the contrast without affecting the resolution. See also Reviewer 1 Question 9 for a similar question on spatial resolution.

3. Table 2 shows that the ratio of prompt/positron emissions varies according to the radio-isotope, from 13% for ^{82}Rb to 100% (^{44}Sc), and ^{124}I and ^{86}Y used here for antibody-labeling having ratios of 51% and 33/83%, respectively. In other words, using ^{124}I alone would yield two times more LORs than VLORs. How is mPET coping with that? Are the annihilation-only LORs rejected? If not, will they not pollute mPET reconstruction and preclude separation from pure beta+ radio-isotopes?

An excellent observation. Prompt-gamma emitters such as ^{124}I will have annihilation-only LORs (i.e. standard double coincidences) as well as triple coincidences. Their relative

amount is known based on these percentages and the attenuation and sensitivity of the scanner (for the energy window used). All these values are estimated with the Monte-Carlo (MC) simulation. The separation algorithm used at the end of the reconstruction process takes into account the difference in isotope abundance. This has been now added in Supplementary Fig. 1 as follows: The separation of both radiotracers is performed considering the relative sensitivity of the detection of double and triple coincidences. This sensitivity is obtained from realistic MC simulations, which take into account all the relevant processes in the emission, transport, and detection of the signal emitted by both radiotracers (pure positron emitter and positron-gamma emitter). In Supplementary Fig. 2a and Supplementary Fig. 3, we provide some examples of event percentages for ^{124}I for doubles, triples, and random triples and show that the double-to-triple ratio is essentially unchanged for the activities used.

4. It is not clear what were energy detection ranges used with the different instruments and how it was optimized. Use precise values and not vague terms such as “can be expanded (line 448)”. How were the various photopeaks matched to account for detection sensitivity at different energy windows?

Thanks for pointing out this deficit. The preclinical PET scanner Inveon used in this work allows selecting the energy window within a user-defined range around the 511 keV peak, being the maximum nominal allowed energy of 814 keV for a LSO detector [PMID17873140]. In this work, we were interested in the detection of the 511keV gamma rays and the 603 keV prompt-gamma from ^{124}I. We used a maximum energy window of 700 keV. This was chosen for several reasons: on one hand, we had some concerns about the data acquired when using the maximum upper energy allowed (maximum number of random triples and noise), and we wanted to avoid the high-energy gamma rays emitted without prompt positron emission from ^{124}I (723 keV, 1510 keV and 1691 keV) [PMC5283715] which may increase the number of random coincidences. A detailed study of the performance of mPET as a function of the energy window selected is out of the scope of this work as energy is not used for reconstruction, but is part of a follow-up study. This could be much better done with PET scanners that provide the energy information of each detected event (such as the Biograph Vision for instance), and with higher upper energy limit, providing more isotope combinations. See also Reviewer 3 Question 2 for a response to the energy window selected.

5. How were the accepted/rejected triple coincidences chosen? Supplementary information lines 148+ should show quantitative data.

The amount of triples in these acquisition is around 4% of the doubles (see for instance the Supplementary Fig. 2a with the sensitivity distribution of doubles and triples). As it is mentioned in the introduction (line 78-79), in most commercial PET scanners, triples are stored in the list-mode data as two consecutive lines-of-response. Therefore, the extraction of the triple coincidences from the list-mode data is based on finding consecutive coincidences with a common detector element. It is important to note that the probability of obtaining with this approach spurious triples (i.e. two unrelated consecutive LORs that share by chance the same detector) is quite low. It depends on the number of detector elements (it can be estimated as 2 divided by the total number of crystals), which in most scanners is below 0.01%. For instance, in the Inveon scanner, the number of crystals is

25,600, and therefore the probability of spurious triples is 0.008%. This can be also estimated directly from the listmode data, analyzing the number of non-consecutive coincidences (for instance separated by 20 coincidences in the recorded data) with a common detector element. We have added these details on line 42-43 in the supplement.

6. The argument about avoiding double CT irradiation is of limited value since most modern PET-CT use low-irradiation CTs. In my view, the synchronization of the images is a much stronger argument, and the authors should rather insist on that point.

We thank the reviewer for this suggestion, and we have toned down the radiation savings from the CT dose, and focused the benefits of mPET on the synchronization of scans as mentioned now on lines 49-51.

7. How will the authors manage the different kinetics variations among patients? Among tumors? Among tumor drug delivery? Isn't the optimization of the injection-imaging delay an unsolvable nightmare?

This is a fascinating question. Patients are indeed heterogeneous as well as their tumors, and in fact, the main advantage of our method with respect to other dual-tracer PET methods is that it does not rely on any kinetic modeling of the patient. Furthermore, clinical workflows would be used as optimized for the radiotracer, thus no additional optimization would be needed when combining protocols for clinical mPET imaging. Our goal was to provide a reconstruction method that would allow the direct in situ comparison of two tracers without the need of known kinetics. Conversely the unbiased reconstruction of mPET would allow scientists and physicians to more quickly optimize a protocol if two mPET compatible radiotracers (from Table 1 and Table 2) could be administered.

8. There is little interest stemming from the nanoparticle drug delivery experiment with mPET. Here the release of the drug from the nanocarrier is completely obvious and could be made evident with other techniques. Isn't there a better example of using slow drug-release nanoparticles?

We agree with the reviewer that this figure on the surface appears obvious, though the ability for mPET to track both nanoparticle and small molecule drug directly is a major improvement over other methods. It is important to point out that our in vitro stability data did not predict this early separation of trametinib from the particle, and we were completely surprised by this result, changing our approach going forward (unrelated to this paper). As such it serves as a great example how an unexpected (and considered "negative") result revealed with mPET can be of significant importance for a project – be it in further animal studies or in patient therapies, Often nanoparticle delivery experiments focus on identification of the nanoparticle alone, or use a dye as the surrogate to the actual molecule intended, citing any statistical increase in intensity despite no direct nor noninvasive proof the drug was delivered by the nanoparticle directly. We didn't explore other radiolabeled drug nanoparticle combinations as it was beyond the scope of this work, and rather we wanted to use dual labelling and reconstruction with mPET to show decisively the failure as a n important and decisive result. Few iodinated small molecules are used mainly due to their long half-life relative to biodistribution, where fluorine

derivatives work well enough, and that dehalogenation of iodine and uptake in the thyroid present as undesired features.

9. Figure 4 and supp. Fig.2: How do you explain the higher FDG uptake under BRAF or MEK therapy?

We thank the reviewer for this great observation. The tumor model used in this experiment is the B16F10 model that had been used in previous studies showing the distribution [^{124}I]I-trametinib, though the B16F10 cell line is a MEK and BRAF wildtype. BRAF sits upstream of MEK. In a paper by Parmenter (PMCID4110245), inhibition of wildtype BRAF leads to an increase in GLUT1 expression in Figure 1h. Alone, MEK or BRAF inhibition should induce this increase in GLUT1, thus also increased FDG uptake, though we were only able to show significance for the trametinib inhibition alone. BRAF inhibition does show elevated ROI values but is not significant. Parmenter et al. also show that upon BRAF inhibition in V600E mutant melanoma lines (not tested in this manuscript), the GLUT1 expression was reduced. This reduction is in line with other accounts with homo and heterozygous BRAF V600E mutations on vemurafenib therapy (PMC3405466), though not for the wildtype melanoma used here, and would be of much interest in future studies. We have included in the manuscript on lines 227-231 these findings to support the observation.

10. Supplementary Fig 3: not the same mice, quantitative values must be provided.

We thank the reviewer for the suggestion, and we have included for each group a tumor ROI analysis as Figure 5f. We did find that there was a barely significant increase in average ROI intensity between the saline and trametinib treated arms of the study. We however did not find a significant difference for FDG in other treated arms relative to saline. The trametinib treated arm clearly has the lowest average intensity as expected for a blocking study, though was not statistically significant.

11. Supplementary Figure 4: it's very hard to read the figures with this color scale. Again, numbers would be necessary for data interpretation.

We apologize for this oversight. The separated images have been calibrated and the scale bars clearly included now in this revision as also requested by other reviewers. If the text is still too small for the figure, we would be happy to increase the font for the tracer organ distribution section, but wanted to keep that minimal to focus on the separated images.

12. Line 126:" expanded energy range" … please give numbers.

We have clarified this line, in accordance with the methods section to describe the increase of the maximum energy range from 650 keV to 700keV on line 149

13. References are not mentioned properly, abbreviations are not standard or absent and some titles (#16) have been omitted.

We apologize for the reference mangling. We have updated the reference style and included the title for reference #16 and appropriate information for reference #4.

14. Please correct typos and English spelling, e.g., foot pad and not food pad, etc. Color scales of the supplementary data are barely readable. Some panel references (Fig. 6 are not indicated in the figures.

We sincerely apologize for these errors. We must have been hungry at the time. We have checked the items above and have made sure spelling and words are correctly used.

Reviewer #3:

This paper innovatively proposed a general multiplexed PET reconstruction method (mPET) to separate the two tracers with single acquisition, relying on the triple coincidences of some specific radioisotope due to the additional prompt gamma rays. To validate the proposed approach, several preclinical/clinical experiments were designed and performed well on phantoms or mice, and the results showed two tracers could be differentiated successfully. The paper is well-organized and clear to read with abundant supportive data and nice images, however, I still have some questions on the described methods and experiments as follows:

Major:

1. From L84 – L104, you gave a brief intro of the new general reconstruction method, and it seems like mostly inspired by the 25th and 29th refs, and explained more in the supplementary files. Considering the algorithm is kind of important for readers to understand the entire method, maybe it is better to re-organize the methods section (e.g. including more mathematical expressions) and present it before the results.

We thank the review for the suggestion to increase the mathematical background of the method to this manuscript. We have added the brief algorithm to the introduction before the results and have provided in the supplement Figures 1-4 more details about the mPET reconstruction performance.

2. For the preclinical phantom experiments, a 350-700 keV energy window was chosen during the acquisition, how was this parameter optimized? Especially, the emitted prompt γ-rays have different energies? Does it require to redesign the PET detectors to detect high-energy gamma rays more efficiently?

We thank the reviewer for these great questions. Review 2 question 4 asked a similar question. Using a uniformity phantom with a known activity concentration of each single isotope, we selected the 350-700 keV range in part to maximize the number of prompt gamma emissions recorded while keeping low the random triple event rate. This range was selected to be used for all Inveon isotope combinations, though in practice the energy range can be easily adjusted as needed in the acquisition setup. We did not find a significant difference decreasing the energy range below 350 keV. However, in a clinical setting, energy windows are fixed, to provide a more robust operation. None the less the clinical mPET example provided in Figure 3 used a fixed energy window and isotope separation was still possible. It is possible in the future that clinical scanners could be slightly redesigned to detect high-energy gamma rays more efficiently and take advantage

of the higher energy range and provide a better mPET acquisition of triple events. See also Reviewer 2 Question 4 for a response about the energy windowing used.

3. In Figure 3, the 2th cavity in C seems kind of weaker, especially at the edges, why caused this? Is it because the recon or other reasons?

We thank the reviewer for this observation. As the abundances of ^{124}I and ^{89}Zr are similar in the tumor phantoms, we have calibrated the separated images and matched the unseparated image best to those intensities. The difference in cavity c at the edges is primarily due to the difference in triples as well as that ^{124}I has a higher positron energy and as such a lower image feature resolution. We have amended the figure description to account for these differences.

4. In Figure 3, The E is standard PET, if the ^{124}I signal could be attenuated as scatter events, then why is the ^{89}Zr region so dark in E and bright in F? Do the E, F and G share the same colorbar?

We apologize for the confusion of image scale. We have calibrated the separated images in Figure 3 to ensure they share the same scale. The newly calibrated images better align the high ^{124}I signal due to the higher concentration in the phantom, and the unseparated image has been adjusted to best match the ^{89}Zr and ^{124}I calibrated intensities.

5. For the second experiment, the [^{124}I]I-trametinib injected 24 hours before the [^{18}F]FDG, considering the ~4 days half-life of ^{124}I, it is reasonable and the results also validated the sampling protocol. Have you ever tried different injection time of the second tracer, or simultaneously? Do they also work well with mPET method?

We thank the reviewer for this question. The timing of the radiotracer injection was based on prior knowledge of each individual radiotracer's biodistribution to achieve the highest uptake in the tumors. We have not tried in this instance dual injection, though we have shown successful quantitative separation in Figures 5, 6 and Supplementary Fig. 14 covering simultaneous radiotracer injection and injections staggered within one hour with great success separating both isotopes.

6. For the nanoparticle drug delivery, do you think the standard PET recon could approximately seem to be the sum of [^{124}I]I-trametinib and [^{89}Zr]Zr-ferumoxytol? In Figure 5, images in the first row seem stronger than the sum of corresponding mPET images in the second and third rows. Anyway, it is better to rescale and make sure the images are comparable, so we could explore more about it.

We thank the reviewer for this observation. It is true that the sum of the separated images should match the standard PET image, but this is only valid before the decay correction of each radiotracer and the use of calibration factors to obtain each separated image in %ID/cc units. We have adjusted the unseparated image in Figure 5 to allow a closer comparison and included the scale bar details which we unfortunately omitted in the first manuscript version.

7. From biological effect point of view, the emitted high-energy γ-rays (more than 511keV) is not desired. On the other hand, PET detectors are designed to detect the pure 511 keV emitter. Therefore, there are some efforts for separating dual-tracer signals from reconstruction point of view. The authors have to address the limitations of the proposed method.

We thank the reviewer for this insight. Indeed, the emitted radiation has known biological effects and the radiation dose in PET is always an important concern. We have included in the introduction now several methods by which the radiation dose could be lowered in future work (on lines 60-66) through artificial intelligence, increasingly sensitive PET scanners, as well as reducing the need for longer lived isotope imaging with pretargeted approaches alongside CT scans for each single PET scan. We have mentioned in the introduction in lines 103-108 that the separation method for mPET relies not on the energy of the emission, but rather the prevalence of the triple event, that is listed as contiguous double coincidences in the listmode. We have included details on how we adjust the energy window of the detector in general, to allow for more triple events, but no energy information is at all recorded. In the future, greater energy discrimination will play a role in higher order multiplexed images, but can utilize the same reconstruction method. We have included some limitations in lines 82-84.

Minor:
1. In Figure 3, A-G is better matched with the annotations a)~f).

We thank the review for this and have changed the annotations to match a-f

2. In Figure 4, A-D is better matched with the annotations a)~d).

We thank the review for this and have changed the annotations to match a-d

3. In Figure 5, the front size of the mice images looks too small, please check it.

We thank the reviewer for this observation. We have shortened the period of imaging shown in the figure to better visualize the text, as no appreciable difference is seen after 24 hours. We have added the original figure as Supplemental Fig. 8.

4. In figure 6, I just noticed the colorbars of 68Ga-PSMA and 124I images have black outlines, maybe the unseparated series should keep the same style.

Thank you for this observation. We have corrected Figure 6 to have the same colorbar style.

Rebuttal 2

We thank the reviewers for directing us to produce a more robust and approachable manuscript on two color PET imaging.

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

I have read the manuscript and the authors' rebuttals. The reviewers' comments were definitely necessary and mostly well-answered by the authors.

I think this paper is getting quite robust, with a few pending issues with:

We thank the reviewer for their comments and helping us produce a robust and accurate paper

1. I still think “Duplexed PET” is more accurate than mPET for this paper. Given what was demonstrated here, it is scientifically more accurate to name this one dPET and the next paper (where they can demonstrate multiple colors or rainbow of colors with better PET scanners) mPET. One needs to learn to walk before they can run.

We understand the need to tone down the title and as such the new title is more directly focused on two color imaging. Furthermore, the manuscript examples we show clearly identify two color imaging is at use here. Unfortunately, recently Sedecal has used the term multiplex for two color PET, so while we could introduce dPET, (and see the reasoning), we feel it would cause more confusion to the field to have dPET and mPET both mean two color PET imaging. We have also discovered that the Inveon system uses dPET as well, deferring to the dedicated PET module, adding additional confusion to readers. We agree that we need to walk before we run, and by toning down main claims of multiplexing and focusing on two color results we believe that we have achieved just that. Thank you.

2. As suggested by reviewer 1, point 5, the authors should look at a case study where the drug-particle combination does work. They didn't do this, nor did they say why they wouldn't do this. Although, I accept the authors' justification for having (and keeping in the final draft) a case study where the combination doesn't work.

We thank the reviewer for this comment. As we have mentioned in the previous revision comments, building radiotracer analogues of drugs is not trivial, and many attempts to add on a radioligand can affect the original drug property, changing nanoparticle loading behavior. As such designing a new radiotracer from a known drug-particle combination may not be in any way a recapitulation of the work and to add resources to prove drug loading works is beyond the scope of this paper. This paper is to highlight the use of dual PET imaging to gain information not attainable

in single PET tracer studies. We believe that the study shows particularly as a “negative result” that the method can provide valuable insight as we were not aware of this rapid separation. Here, not the nanoparticle drug loading is the result but the fact that mPET can reveal unexpected results

3. The authors' treatment of Figure 5 and Supplementary Figure 8 is still not great. Max and min % are still not stated and a comparison between the standard PET and mPET (should be dPET) with different scale bars is not straightforward as the figure may imply at first glance.

We thank the reviewer for the insistence on making each figure as quantitative as possible. We can show the unseparated images calibrated to one or the other isotope, but by introducing a triple emitting isotope, standard reconstruction does not reconstruct these triples, thus an even calibrated image by any estimate is inaccurate. This is partly why triple emitting isotopes have been considered inferior to pure emitting isotopes like ^{11}C and ^{18}F. We provide the Min-Max range to represent as close to the combination of each separated image. We have reviewed Figure 5 and have adjusted some of the initial timepoint images to better match the separated images.

The responses to much of the comments were discussion/arguments, albeit politely. Other authors may have done significantly more experiments or made more concrete changes to the manuscript.

We thank the reviewer for their persistence and do appreciate it. We added some more data as indicated above. However, particularly the requested case study where the drug-particle combination works is not a trivial one and would, as indicated, require considerable resources and time for synthesizing a new radiolabeled drug, and load it into the particle, followed by in vitro characterization and then animal studies - with no guarantee that this would provide the preferred “positive result”. And at the end, even if positive, the study would only have provided the same result – that mPET can monitor these therapies, no matter whether it is a positive or negative result. Therefore, we appreciate the reviewer’s acceptance of our discussion.

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

My opinion is that the authors addressed all of my comments and made the changes and additions accordingly.

We thank the reviewer for their direction in making this manuscript better.

Minor changes to improve the manuscript:

`line 158, a word is missing

We have amended this sentence as such “The full energy range of was not selected as to reduce random event noise and multiple coincidences from Iodine-124 (^{124}I)”

line 123 please define Wv1 and Wv2 in the text.

We thank the reviewer for the clarification, we have added to equation line 1 that Wv means isotope weight by abundance.

line 505 It would be nice to comment on Table 2 by indicating which prompt-emitting isotopes could be used for mPET on any PET scanner without modifying the instrument, and which would require improved energy resolution, etc. The reader interested in the new technology could take a glimpse at the kind of double tracer experiments he could plan with respect to the domain of imaging (cancer, brain, cardiovascular...

We thank the reviewer for this suggestion. We’ve added a column to Table 2 to state what prompt gamma isotopes would be visible, even partly in the 350-700 keV energy range used. As PET detectors have different energy resolutions we’ve conservatively given a few options that are feasible. Expanded energy ranges for example to 814 keV would have enabled at least two more isotopes.

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

In this revised version, the authors have addressed my concerns.

We thank reviewer 3 for improving the quality of the manuscript.