We would like to thank Reviewers for carefully reading our manuscript and the opportunity to address all concerns raised, improving greatly our manuscript.

Reviewers' comments:
Reviewer #1:
Manuscript Summary:
The manuscript entitled "Generation of zebrafish larval xenografts and tumor behavior analysis" submitted by Rita Fior and collaborators describes step by step how to generate xenografts in zebrafish embryos and gives useful guidelines for tumors studies. This manuscript is well constructed and provides a useful protocol for researchers working on cancer-related topics who wish to use the zebrafish model.

Minor Concerns:
Before accepting this manuscript for its publication in JoVE journal, the authors need to make some adjustments cited below:
- From the line 312, the authors stated "to position the Eppendorf tubes horizontally". I don't know if this position is necessary since there is 1.3 mL in 1.5 mL Eppendorf tube and the zebrafish embryos are in the bottom of the tube.

We thank reviewer 1 for this comment- sorry for not being clear. We recommend placing tubes horizontally, exactly to avoid the accumulation of embryos in the bottom. This prevents the xenografts from being homogeneously fixed, or permeabilized with the solutions. To address this, we recommend placing tubes horizontally to allow more space between the xenografts which in turn improves contact between the fixative and permeabilization solutions. We now introduce a small note to explain this recommendation.

- Line 410, it is not clear why, to obtain the total number of cells in the tumor, the authors use 1.5 as a dividing factor in the formula. This point needs explanation.

We thank the reviewer for raising this concern. We have included more information in the protocol as well as in the discussion section.

Our set up to acquire confocal images is z stacks with 5µm interval and our control human tumor cells (HCT116) have an average nuclear size of 10–12µm of diameter. This means that if we count the DAPI cells in slice 1 and then slice 2 ~50% of the cells are shared between both slices in slice 2. Therefore, if a researcher counts every slice – they need to be very careful and in every other slice have to be always double checking if they are not counting twice the same cells. This is not only very time consuming but also error prone between researchers and even with the same researcher.

We came up with a compromise (Fior et al 2017): count the first, middle and last slice and then do the correction (if we multiply by the whole stack, we would be counting in every other slice more 50%). We did some tests:

- The average error of manual counting of the whole tumor between researchers was ~20%.
- But 2 researchers using the formula had a ~2% error.

The use of this formula has a 93% accuracy and 98% reproducibility rate. We also tested automated methods, but they demonstrated an error higher than 50% due to threshold settings.However, this is the formula for cells with an average nuclear size of 10–12µm of diameter – if cells are bigger or smaller the correction must be adjusted.

- In this protocol, calibration of the number of injected cells is missing. They must use a "calibration slide" and mineral oil to quantity and adjust the volume injected from the needle.

We thank reviewer for raising this point – however we do not calibrate the injection needles. We have now addressed this concern in our discussion.
Contrary to the delivery of oligonucleotides or drugs into the zebrafish embryo, we do not use graticule to calibrate the injection needle when working with tumor cells for xenografts. The concentration of cells changes along the capillary, in the beginning cells are more diluted but then become more and more concentrated until they start clogging the needle. Then it is necessary to cut the tip of the microinjection needle to increase its diameter or change the needle altogether. These problems hamper graticule calibration - because 5 min after calibration the original calibration is gone.
To address this issue, the number of cells dispensed is regulated by eject pressure and time needed to reach a size similar to the embryo eye within 1-3 pulses. Then, to further control the tumor size, at 1dpi, xenografts are sorted according to the tumor size as shown in **Figure 6 B-B"**. As represented in **Figure 6C** example, this method of sorting is efficient in reducing the variation of tumor sizes: if we pool them all together (+,++,+++) the STEV is ~double of the ++class (~906 cells to ~422 cells) and the coefficient variation is ~31.9% compared to 14,5% in the ++class. Since the reference for injection is volume, the total number of cells varies a lot between cell types – being dependent on the size and shape of the cells. For example, big cells with a lot of cytoplasm like breast cancer Hs578T, produce much smaller tumors (~600cells). Also, each cell line requires different number of cells. For instance, the HT29 CRC and RT112 urinary bladder cancer cell lines have shown that the higher the number of cells injected, the higher the zebrafish mortality. Therefore, a period of optimization while developing the xenografts is needed to test if the cell line has toxic effects in the embryo or requires higher/lower densities of injection.

Reviewer #2:

Manuscript Summary:

The goal of the protocol is to engraft cancer cells into zebrafish embryo's to form tumour.The PDX in zebrafish are then used test therapy regimens and score for tumour response. Every part of the protocol is very detailed; engraftment, therapy application and read-out. The writers also share many tips and tricks to overcome hidden hurdles. The research group clearly has much expierence seeing the detailed list of cell lines that they engrafted already. I believe that with little expierence of zebrafish handling, one should be able to execute the protocol.
The advice that practicing is necessary is really relevant, since the executer can't expect good results the first couple of times.

Major Concerns:

There is nothing mentioned about zebrafish xenografts dying during the experiment and how this is handeled during analysis.

We thank reviewer#2 for this comment- we now included information addressing this concern in the protocol. Any dead or swollen xenografts are discarded, and these xenografts are not fixed and not considered for quantification. Drug treatments and some tumor cells can induce toxicity and eventually cause xenograft mortality. Concerning drug treatment, since before any treatment we find the maximum tolerated concentration (maximum drug concentration without inducing mortality i.e. max 5%) we do not encounter this problem often- therefore we simply discard these xenografts. However, when the xenografts die because of toxicity of the cells – then mortality should be quantified along experiments – since they can be an interesting phenotype to pursue.

We also changed our formula to specifically state that the xenografts used for quantification are live xenografts.

Minor Concerns:

Some tools to use in the protocol such as agarose plate and hairpin are visualized but not really explained how to make it.
A picture protocol on how to do this, would be helpful

We thank reviewer#2 for the suggestion and have now expanded the information regarding agarose plate preparation and hairpin assembling in our protocol as well as a picture protocol (new Figure 2D).

The readout is very well explained until the percentage of markers.
It says: "Quantify manually the positive cells along the stack with the Cell count plug in."Are all positive cells counted, do you take 3 z-stacks like for tumor size calculation, do you take a treshold of which cells are positive? What with merged/ overlapping cells?
A step by step example of result analysis should be helpful.

We thank reviewer#2 for the pointing this out- we have now included a step-by-step guide for datasets quantification.

- Tumor size:

Our set up to acquire confocal images is z stacks with 5µm interval and our control human tumor cells (HCT116) have an average nuclear size of 10–12µm of diameter. This means that if we count the DAPI cells in slice 1 and then slice 2 ~50% of the cells are shared between both slices in slice 2. Therefore, if a researcher counts every slice – they need to be very careful and in every other slice have to be always double checking if they are not counting twice the same cells. This is not only very time consuming but also error prone between researchers and even the same researcher.

We came up with a compromise (Fior et al 2017): count the first, middle and last slice and then do the correction (if we multiply by the whole stack, we would be counting in every other slice more 50%). We did some tests:

- The average error of manual counting of the whole tumor between researchers was 20%.
- The same researcher counting every slice carefully vs the formula had a 3% error.
- But 2 researchers using the formula had a 2% error.

The use of this formula has a 93% accuracy rate and 98% reproducibility rate. We also tested automated methods, but they demonstrated an error higher than 50% due to threshold settings.
However, this is the formula for cells with an average nuclear size of 10–12µm of diameter – if cells are bigger or smaller the correction must be adjusted.

To quantify activated caspase3 is trickier – to define what is positive – if it is one or 2 positive cells. We recommend that the same researcher has to count control and experiment i.e. the errors made in control will be also performed in experiment. Also, a new researcher needs to count old images that were already quantified by more experienced researchers to calibrate their eyes. We tried several other automated methods. However, in our hands they generated more error: the same image, the same researcher, 5min after would give a totally different result.