Author’s response to reviews

Title: Assessment of anti-inflammatory tumor treatment efficacy by longitudinal monitoring employing sonographic micro morphology in a preclinical mouse model.

Authors:

Sanjay Tiwari (stiwari@email.uni-kiel.de)
Jan H Egberts (Jan-Hendrik.Egberts@uksh-kiel.de)
Olena Korniienko (lena_kornienko@email.uni-kiel.de)
Anna Trauzold (atrauzold@email.uni-kiel.de)
Claus C Glüer (glueer@rad.uni-kiel.de)
Holgar Kalthoff (hkalthoff@email.uni-kiel.de)

Version: 3 Date: 3 May 2011

Author's response to reviews: see over
Response to Reviewer 2 (James Lacefield)

Major Compulsory Revisions

1. [Results, 2nd paragraph and Fig. 1A]: It is not clear whether the observed difference in growth rate of the two groups from Day 7 to 12 is a meaningful result because the tumor volumes on both days are not significantly different, so the apparent difference in slope may simply be the result of volume measurement variability. The fact that the mean control tumor volume was less than the mean treated tumor volume on Day 7 in particular gives the impression that measurement variability was a factor here. Is a difference in slope still observed if the change in volume of each individual tumor is computed and then those volume changes are averaged, as opposed to taking the difference of averaged volumes?

   In the results section, 2nd paragraph we stated that “the growth rate revealed no significant difference from day 7 to day 12”. Perhaps the reviewer is referring to the next time window from day 12 to 17? We reported a significant difference in growth rate from day 12 to day 17, although the tumor volumes for both days are not significantly different. As correctly stated by the reviewer in principle there are two different ways to calculate the difference in growth rates:
   - If the change in volume of each individual tumor is computed and then those volume changes are averaged we obtain the result stated in the paper: $14.0 \pm 3.2 \text{ mm}^3/\text{day} \text{ vs } 24.7 \pm 3.8 \text{ mm}^3/\text{day}$, $p<0.05$ (assuming equal variances for both groups which was tested to be true). This approach maximizes statistical power exploiting the ability to longitudinally measure rates of chances in individual animals. This is precisely the advantage of non-invasive longitudinal sonographical imaging which permits to reduce the interindividual variability by comparing changes measured within each animal across groups of animals, e.g. in different treatment groups.
   - If we first average the group means for each day of measurements, and then test whether there is a group difference for the difference of the day measurement time point we have the same average values of 14.0 vs 24.7 but this difference shown only a trend: ANOVA of volume versus group and day yields $p<0.093$ for the interaction term of group x day.

   In conclusion, by using the more powerful statistical approach we can demonstrate a statistically significant difference in tumor growth rates between days 12 and 17. The chances that this difference is due to random measurement variability is below 5% (type I error). The magnitude of the difference between the groups is limited reaching physiologically relevant levels only during the third time window from day 17 to 21.

2. [Results, 2nd paragraph and Fig. 1B]: What is the likely explanation for the fact that the variability of the control tumor volume was greater than the variability of the treated tumor volume, but the variability of the treated tumor mass was greater than the variability of the control tumor mass? This apparent inconsistency was not observed in the second (resected model) study (Fig. 2B).

   The reviewer notes that the variability of the treated tumor mass was greater than the control tumor mass, but by ultrasound the volume of the control tumor had greater variability than the treated tumor. We can see three reasons for this:

   - First, the measurement errors of sonographically determined tumor volume and physically measured tumor mass after resection are different. For the latter
method smaller tumors in the orthotopic setting are generally more difficult to resect and measure leaving more room for variability. For ultrasound on the other hand larger tumors typically extend deeper into the tissue and thus we have to deal with decaying signal intensity and reduces contrast to noise ratios. This makes it more difficult to determine the borders of the tumor tissue on the side opposite to the surface contributing to increased variability.

Second, this may be partly simply a statistical artifact. In the control group we have only 7 animals and thus a single data point can have a strong impact on the box plot shown in figure 1B. Indeed if we deleted only one extreme case from the control group the interquantile range (10%ile to 90%ile) is almost cut in half (being reduced from 605 mm$^3$ to 329 mm$^3$), now considerably closer to the interquantile range of the infliximab group of 197 mm$^3$. Of course it is not acceptable to exclude such an outlier and thus this was not done in the paper but perhaps this example demonstrates the limits of the statistics with a limited sample size.

Third, we have also addressed this issue retrospectively by reanalyzing ultrasound data and performing histochemistry. We do not attribute this apparent inconsistency to inter-operator variability since only one operator performed the ultrasound imaging and analyses. Furthermore, the hypoechoic contours of the tumor were largely distinct in both primary and recurrent tumor. However, dissection of the tumor post mortem was performed by two operators and inter-operator variability in defining the tumor boundary for dissection may have had a role. Upon dissection it was noted that the control tumor was a more defined solid lesion than the primary treated tumor. We have now performed immunohistochemistry of control primary tumor and primary treated tumor and notice that the control primary tumor has a more dense tumor parenchyma than the treated tumor. The cellular composition of the stroma of the treated tumor is as of yet uncertain but may contain more fibroblastic and fatty infiltrates. If so, this may account for the variability observed in measurement of tumor mass. In view of this finding we have included a sentence in the results section, 2nd paragraph stating 'Immunohistochemical detection of cytokerratin was performed on cryosections of excised pancreatic tumor tissue. Control tumors appeared to have a greater parenchymal tumor density than Infliximab treated tumor (Figure S1). The cellular composition of the stroma is currently under investigation'. We have attached a figure as supplementary data depicting 4 representative staining of the primary control tumor and 4 primary Infliximab treated tumor to demonstrate the difference in cellular composition. Interestingly, the observation of fatty infiltration in primary pancreatic tumor following neoadjuvant therapy is mentioned in the literature and warrents further investigation (Makay, O; Kazimi, M et al., Fat replacement of the malignant pancreatic tissue after neoadjuvant therapy. Int J Clin Oncol (2010) 15:88–92).
A. primary control-1 primary control-2 primary control-3 primary control-4

B. primary infliximab-1 primary infliximab-2 primary infliximab-3 primary infliximab-4
both animal models and the greater statistical significance of the difference in end-point mean tumor volume compared to end-point mean tumor mass are the basis for stating longitudinal ultrasound measurements to be "more reliable" for monitoring therapeutic efficacy. Perhaps more important is the ability of noninvasive sonography to examine the dynamics of tumor growth longitudinally in a given animal which provides insight into the temporal development of therapeutic effects. However, we take the reviewers point that “more reliable” maybe an inappropriate phrase since it implies one method was robustly tested against the other (i.e. experimental associated errors of large vs small tumors, inter operator errors etc). We would like to rephrase the previous three sentences to be more fitting to the issues raised by the reviewer and we have added a fourth sentence covering the topic of dynamics which also nicely bridges to the next paragraph of the discussion section:

‘We report here that longitudinal sonography measurement of tumor volume in vivo confirmed the inhibitory effect of Infliximab on primary tumor growth in both the resected and the non-resected models. End point differences in volume measured by sonography showed agreement with differences in tumor weight between the two groups measured post mortem. The findings indicate that for both primary and resected tumor models of pancreatic cancer, longitudinal ultrasound measurement can be used to reliably monitor therapeutic efficacy in vivo. Additionally, it permits to study the temporal development of tumor growth sequentially in a given animal providing more detailed insight into the dynamics of tumor growth and therapeutic effects.

4. [Discussion, 2nd paragraph]: It is not clear that these studies actually demonstrate detection of "complex" tumor growth kinetics. Figs. 1A, 2A, and 3 appear to show exponential or even linear tumor growth; it is debatable whether any of the longitudinal data sets qualifies as complex.

We observe different kinetics of effect of drug activity in the two models, the first model acting later and maintaining its effect while the second model acts earlier and appears to have a transient effect. In the first model, Infliximab treatment reduces the growth to a linear increase in contrast to an exponential increase seen in the untreated tumor. In the second model, untreated tumor increases at a linear rate with the treated tumor increasing at a slower linear rate. Indeed it is debatable whether this qualifies as complex pattern of growth since we don’t observe variation in responses at different time points including zero growth, regression and/or subsequent resumption to exponential growth. In view of this we have modified the sentence to state our observations in this experiment:

‘Longitudinal sonography thus allow to detect variable rates of tumor growth not accessible by post mortem measurements of tumor volume or weight.’
Minor Essential Revisions

1. [Methods, 4th paragraph, minor issue not for publication]: The axial spatial resolution of the 40- and 25-MHz transducers is 40 and 70 microns (not mm), respectively.
   Changed in manuscript.

2. [Methods, 5th paragraph]: Please specify the spacing between image planes that were segmented to estimate tumor volume as this is a user-selected parameter in the VisualSonics software that can affect the accuracy and variable of the volume estimates.
   The following sentence was added to the methods section:
   'Two dimensional images were initially acquired at regular spatial intervals which were parallel and uniformly spaced at 100µm.'

3. [Results, 1st paragraph and elsewhere, minor issue not for publication]: The precision to which most of the p values are reported suggests that these are the actual p values, not the threshold of statistical significance, so these data should be reported as, e.g., p = 0.016, not p < 0.016.
   Changed

Discretionary Revisions

1. [Results, 5th paragraph]: Consider including a scatter plot with the best-fit regression line to illustrate the comparison of ultrasound tumor volume and end-point tumor mass. Why was a volume-to-mass correlation coefficient not reported for the first study?
   The correlation of sonographically measured volume and physical measurement of weight was borderline significant (r² = 0.2, p = 0.066) and was strongly influenced by one outlier (after exclusion of this outlier r² = 0.31, p = 0.02).
   We have added a corresponding statement to the results section. We did not add a figure since we felt that this would not provide substantial additional valuable information for the reader beyond statement of the level of the coefficient of determination.

2. [Results, 6th paragraph]: Sample images showing the ultrasonic appearance of a representative tumor and surrounding pancreatic tissue would increase the usefulness of the paper to other researchers who are considering using ultrasound to study mouse pancreatic tumor models by helping those readers understand what to look for in their images. Sample images of the peritoneal tumors might be the most interesting as this seems to be the most significant advantage of ultrasound over conventional end-point analysis in this study.
   We will now include an additional figure in the methods section with ultrasound images of the primary and peritoneal tumor.
Figure 1. Representative images of pancreatic tumor margins by ultrasound imaging. Difference in echogenity define the tumor from the surrounding tissue. Tumors are hypoechogen (dark-grayish) compared to surrounding tissue. (A) primary orthotopic pancreatic tumor (red arrows). The light coloured specks (hyperechogenic) within the tumor may be microcalcifications or fatty deposits. (B) The cross-sectional area of the primary pancreatic tumor is depicted. The black areas (anechoic) within the tumor are fluid filled cysts correlating to necrotic areas. (C) Peritoneal tumor with surrounding skin.