Author’s response to reviews

Title: In vivo MRI volumetric measurement of prostate regression and growth in mice

Authors:

Kent L Nastiuk (knasituk@uci.edu)
Hui Liu (hliu3@uci.edu)
Mark Hamamura (Markjham@uci.edu)
L. Tugan Muftuler (muftuler@uci.edu)
Orhan Nalcioglu ( nalci@uci.edu)
John J. Krolewski (jkrolews@uci.edu)

Version: 2 Date: 7 May 2007

Author’s response to reviews: see over
Dr Annabel Phillips  
Senior Assistant Editor  
BMC-series Journals  
BioMed Central  
e-mail: editorial@biomedcentral.com  

Drs. Phillips, Roberson, Heverhagen, and Waterton  

Thank you for the thoughtful review of our manuscript, MS: 3179008001274344, "In vivo MRI volumetric measurement of prostate regression and growth in mice". Enclosed is our revised manuscript incorporating the changes in response to critiques, with the individual responses to the criticism detailed on the following pages.  

Sincerely,  
Kent Nastiuk  
John Krolewski
Major Compulsory Revisions:  None

Minor Essential Revisions:
1. Reference to Figure 1F changed in text to 1E
2. Discussion line 16 extraneous "the" removed

Discretionary Revisions:
1. The initial MRI sequence used created images with chemical shift of fat layers. The image shifting appears to result in an apparent altering of the prostate contours (p 10 lines 5-6). There is a tendency to want to contour the prostate allowing for the shift of the fat tissues, at least a best guess at the effect. It was mentioned that the CHESS images were more reliable, but the point that the non-CHESS images produce less than accurate results might be discussed more thoroughly. For instance, how much would it have affected the regression of the prostate given that each time point could have a different impact of the chemical shift of the fat layers?

It is indeed an interesting idea for discussion. Unfortunately (as Dr. Roberson correctly suggests), the fat surrounding the prostate appears to become more prevalent with increasing time from castration, making any attempt at comparison between the two methods quite complicated. A great deal of additional data would need to be collected (at great expense to us) using the non-CHESS technique, data which we would have no other use for as we suspect that it is of inferior quality. We therefore must defer such discussion to a more qualified investigator with an interest in this question.

We have added the following to the discussion of Figure 3 in the results: Final sentence of paragraph 1 of CHESS section: This fat signal complicates the drawing of the border of the prostate as some estimation of the underlying margin is required. Next paragraph, sentence 4: In addition, The intense abdominal fat (labeled F in Figure 3A’) and other fat signals have been eliminated, and the prostate borders are well delineated (Figure 3B’), greatly facilitating precise quantitation.

2. The CHESS technique being used as a follow-on work to the original affects the read of the abstract results. On first read, the third sentence presents results of prostate regression studies, then the fourth sentence goes back and talks about the CHESS method. The abstract should be improved to differentiate the CHESS technique in the methods section as the follow-on technique used to improve the prostate volume determination and differentiate the sub-volumes of the prostate.

The abstract has been extensively rewritten to differentiate the methods and subsequent results.
Dr. Waterton:

**Major Compulsory Revisions:**

1. This needs far more detail since the use of aged mice is novel. We are told some are ex-breeders and some are >1yo. It seems that 10 mice were used but this should be stated. Please consider providing a table of "mouse demographics" including precise genetic background, supplier, age and whether or not ex-breeder

All mice were retired breeders more than a year old and the genetic background is already included in the methods. This has been clarified in the methods section. Data from 14 mice are presented in this report (additional mice were used in the preliminary studies to establish MRI/anesthetic conditions and validate anatomical identification, see DR#7, below).

2. Please describe the condition of the mice following the repeated bouts of anaesthesia and surgery. Were there any adverse events. What was the pattern of weight loss/gain?

Mice did receive repeated anesthesia, however once optimized, there were no adverse events. Surgery (the castration) was performed only once. The text was modified to clarify this point. In addition, the following sentence has been added to the methods: **There was no weight gain nor loss greater than 10% for any animal during these manipulations.**

3. I infer that there are three substudies: a reproducibility study involving five mice, an involution study involving 3 mice and a regrowth study involving 2 mice. This should be described in detail in "Methods". In particular for the reproducibility study it should be stated how many repeat scans were done for each mouse, and at what intervals.

The number of imaging sessions for each mouse presented for the reproducibility study is detailed in the legend to Figure 1, and the following sentence added to the methods: **We successfully imaged the same mouse up to 3 times over a period of 14 days to assess normal prostate volume determination reproducibility, and subsequently imaged the same mouse up to 8 times over a period of 40 days (b7m1) and 10 times over 37 days (b5m2) to follow castration-induced regression and DHT-induced regrowth of the prostate.**

4. Segmentation: It is well-known that segmentation is a major source of systematic and random error in MRI determination of tiny morphologic change. Much more detail is required. How many segmenters were used? How were they trained? Was segmentation performed over a shorter or prolonged period of time? Were the segmenter(s) blinded to time-point? to animal? What was the intra-scan intra-segmenter reproducibility? If >1 segmenter, what was the intra-scan inter-segmenter reproducibility?
While these are certainly important considerations, this level of image analysis seems more appropriate for patient images, rather than a small study of mouse physiology. As the reviewer points out, there is no other similar study in the literature, other than one image set in Fricke, et al. The primary author is the only expert at this particular quantitation, and hence is self-taught. The primary author performed the segmentation within days of the image capture (not blinded), and the initial set of images was reviewed by a general mouse pathologist, who concurred with our segmentation. Additionally, as is noted in the methods, several mice were examined post-MRI (by both the primary author and the mouse pathologist) to confirm the anatomic localizations. The following sentence has been added to the methods:

**Volumes were determined by a single segmenter at least twice for each image session and agreed within 5%, or they were redetermined.**

5. Volumetry – this is confusing. Under "Methods-MRI" we are told that the 2DMSME provided the images for volume determination. If so, why give the CHESS voxel volume in addition to the 2DMSME voxel volume under "Methods-volume determination"? In Figure 3/4 it CHESS volumes are being used. We are told that volumes were determined twice. Is this by the same segmenter? In the same segmentation session? blinded? both from the 2DMSME or from CHESS?

Both the initial method's voxel size and the improved CHESS voxel size (due to decreased slice thickness are given to show the improvement achieved. The following sentence has been modified in the methods: **The voxel size for the initial (non-CHESS) studies was 0.0041 mm$^3$, and for the subsequent CHESS images it is 0.0027 mm$^3$.** Segmenting is covered in point 4, above.

6. Reproducibility. The statistical treatment is inadequate. The reproducibility in each mouse can be assessed from the test–retest standard deviation, or if the data are lognormal, the test–retest coefficient of variation (CoV). For each subject, $s$, the CoV is the standard deviation, $\sigma_s$, for the repeat measurements on that subject, divided by the mean volume, $\mu_s$, for the subject. Even when as few as two or three subjects is used then $\sigma_s$ is an unbiased, albeit imprecise, estimate of the standard deviation provided Bessel's correction is used. The overall test–retest CoV for a group of $N$ subjects is then

7. "Reproducibility was very good" - this is a judgement which belongs in "Discussion", not "Results". Please give the actual values of reproducibility.

The results have been changed to include the actual reproducibility according to the tests the reviewer thoughtfully included.
8. A major study limitation is that no sham or vehicle group was included, so we cannot really say to what extent the changes result from the intervention, or whether they are confounded by study procedures such as surgery or repeated anaesthesia. In addition N=2-3 is too small for robust significance testing. These limitations should be discussed objectively.

The literature using histological and gross methods over the past 50 years exploring this issue indicates that sham catrated rodents do not undergo prostate regression.

The primary purpose of this investigation is to radically reduce the number of animals that need be used to monitor/model androgen-withdrawal induced regression of the prostate. Thus, while two animals is not sufficient to produce good statistics, given the magnitude of the response, it is adequate to demonstrate the utility of the method and let the reader decide if it accurately reproduces the results in the historical record. The following sentence has been added to the discussion: Only five animals were used in the regression and regrowth studies, whereas a minimum of 4 animals/timepoint at 10 timepoints, 40 animals total, would be required to produce similar data histologically.

Minor Essential Revisions:

1. Abstract implies that all volumetry used CHESS
   The Abstract has been modified.

2. Actual reproducibility should be stated – do not use "very"
   The reproducibility has been reported in the results section, and the conclusion of the abstract modified.

3. Accuracy was not measured – use "precise"
   This has been changed appropriately in the results section.

4. "up to 10 imaging sessions over 5 weeks" – this is not clearly described in Methods

5. The first DNS paragraph seems to summarise the outcome of a previous pilot study. If this is the case, please make that clear, and ensure that the results of the three substudies (corresponding to "Methods") are kept separate from this summary of pilot studies.
   The first paragraph of the results section has been moved to the methods section and labeled as pilot studies. Additional details of the extended imaging are included there.

6. Some of the text in the second paragraph belongs in "Methods"
   This text has been moved.
7. The average volume across 7 animals was 23.8 +/- 0.98. I assume this is SE: surely SD is the more appropriate statistic to describe inter-animal variability.

The results have been corrected to indicate the SD.

8. The voxel size for CHESS of 0.00274 differs from that given under "Methods".

This number in the results has been rounded to 0.0027, as it is in the methods.

9. There are no striped bars in my Fig 3.

The text has been changed to indicate 'grey bars'.

10. In the first paragraph "accurate/accurately" is inappropriate, as no "gold standard" comparator was employed: please use "precise/precisely".

This has been changed.

Discretionary Revisions:

1. The background is adequately described. Some, but not all, previous MRI literature on volumetry, reproducibility and response to intervention in normal prostate and orthotopic prostate cancer is cited. This reviewer is aware of prior studies: o in genetically modified mice (e.g. Degrassi 2007; Fricke 2006; Garcia 2006; Shukla 2005; Gupta 2004/2001; Abdulkadir 2001; Eng 1999; Hsu 1998)

o in other species including rat, canine and primate

The authors may wish to consider a more complete summary of prior studies in the mouse, and possibly in other species.

We are aware of the above mouse studies, and have included some of them in our background, but in the interest of brevity have not been comprehensive in citing all of the literature relevant to tumor volumetry, as it is distinct from volumetry of normal prostate, as we feel this is best left to someone more expert in the field of imaging, such as the reviewer.

4. This referee is not aware of significant MRI literature (with the exception of Fricke 2006) on volumetry, reproducibility or response to intervention in normal prostate or orthotopic prostate cancer in non-transgenic mice, and in particular there seem to be no substantial studies reported on aged mice. The authors may wish to distinguish their study, in the abstract, title, introduction and/or discussion, by emphasising that they are characterising this aged mouse model.
While this is an interesting idea, we are unaware of any literature indicating any
difference between aged and young mouse prostates, and do not present any data on
young mice, so feel it unwarranted to comment.

3. Please consider adding a little information about the DNS pilot study
described in the first paragraph of "results" e.g. total number of mice used in pilot study
Unfortunately, this is quite complicated as there were a variety of failures of both
equipment and anesthesia as we learned how to attain images of sufficient quality for our
analysis. In total, we used seven mice, including the two described in the methods as
used for anatomic comparison. We see no utility in reporting this information for other
researchers to reproduce our protocol.

4. We are told that procedures were approved by UCI IACUC but nothing for
example about humane endpoints. The authors are invited to review the FRAME recommendations
http://www.frame.org.uk/reductioncommittee/journalguidelines.htm for reporting studies in living animals, and to consider whether adopting any of
those recommendations would benefit the manuscript

The major FRAME recommendations are required for approval by the UCI IACUC committee.

5. Point estimates of CoV are imprecise unless very large numbers of mice
are sampled: in the case of a small sample, it is also helpful to quote the
upper bound of the 80% confidence interval on the CoV:

The 80% CI has been added to the results for the estimation of the normal prostate volume.

6. The N in the involution and regrowth studies is too small for formal
statistical analysis, although it might be helpful to report which of the changes exceed 1.96xCoV (or more conservatively 1.96xupper_bound).

All of the involution and regrowth studies were done using a single imaging session at
each timepoint, so no CoV can be calculated for the individual mouse prostate timepoint
measurements, and each value is presented in Figures 2, 4, and 5.

7. Please indicate as far as possible whether there were any differences in
volume with supplier, age, or breeder status.

Unfortunately, neither vendor supplied exact ages for the animals. All were designated
retired breeders, and there was no apparent relationship between volume and vendor.
Major Compulsory Revisions

The entire manuscript is disorganized. Many descriptions of the materials and methods are stated in the results. Also, other parts of the text are in the wrong section. This makes it hard to comprehend the manuscript, and needs to be organized better, e.g. the entire first paragraph of the results section belongs into materials and methods.

The manuscript has been reorganized to address this comment. Please see responses to other reviewers.

The authors need to describe their methods more thoroughly: How many animals did they study in total? What was exactly done with how many mice? How many animals were studied multiple times? What was the exact acquisition protocol? Which sequences were acquired in which animal? In the results section, the authors describe the use of T1-weighted sequences. This is not described in methods.

The methods section has been extensively modified. The data from each animal imaged has been presented, apart from the preliminary studies, which were moved at the reviewers suggestion to the methods section, rather than the results. The exact acquisition protocols are described in the methods for the data presented. The T1-weighted sequences are not described since the results are not presented.

Was there a 90% or a 100% survival rate? 3.5 hours of anesthesia seem to be too long. Also, for a necessary high throughput this acquisition time is much too long and not practicable.

The survival rates have been clarified in the text. 3.5 hours of anesthesia is quite long, which is why we performed the preliminary studies to optimize survival. We are not suggesting anyone use this technique for high-throughput studies, and in fact suggest that using multiple imaging sessions on the same mouse eliminates the need to do such studies.

Please provide the results from the comparison to histology.

Histological changes have been extensively documented in the literature. We compare our data to histological data in the discussion section.

Chemical shift is a well known artifact. Therefore, it description especially in the results and discussion sections is too extensive. Please shorten it substantially.
While we agree that CHESS is well known to experts, such as the reviewer, our intended readership includes many biologists without imaging experience. The two sentences in the results (we find none in the discussion) which describe CHESS allow the non-spectroscopist reader to appreciate how the technique functions.

The authors describe that the higher resolution sequence provides better results. However, what was the gold standard that validates this statement?

We address this issue below.

What was the dose of the DHT administered daily? How much was that compared to natural blood levels of androgens?

The dose is described in the methods section. It is the same dose that has been historically used in the mouse prostate literature.

The entire discussion needs to be reorganized.

Please indicate exactly how the reviewer would like the "entire discussion" to be reorganized, as we believe it to be well organized.

-------------------------------------------------------------------------------

Minor Essential Revisions

The authors use the terms dorso-lateral, dorsal and lateral lobe interchangeably. They should clarify it, and use it consistently.

This has been corrected in the manuscript.

The methods section of the abstract is too short, and needs to be expanded.

The methods section has been expanded.

In the background section, the authors state that the resolution of the references 5 and 6 was not good enough. However, for an organ that is about 20 mm$^3$, a resolution of 0.39 and better should be enough to get a good estimate. Please rephrase.

We did not state that the resolutions obtained by others "was not good enough", only that our studies achieve a higher resolution. We would argue that a higher resolution should provide a more precise estimate of the volumes. A voxel at 0.39 mm$^3$ represents 1/50th of the total volume of the mouse prostate, while our 0.0027 mm$^3$ voxel represents ~1/7500th of the volume.

How was the effect of chemopreventive drugs assessed with only one time point (refs 8 and 9)?

The sentence has been rephrased as follows: *The effect of chemopreventive drugs*
on the volume of the TRAMP mouse prostate was assessed at a single post-manipulation timepoint by MRI[8, 9].

The last paragraph of the background should state the purpose of the manuscript. Currently, it provides a conclusion rather than a purpose. Please rephrase.

The last paragraph has been rephrased as follows: **Our goals in this study are to improve the reproducibility, precision, and resolution of mouse prostate volume determination by MRI. Additionally, by developing a technique to allow imaging the same animal over time during hormone level manipulations, we will be able to greatly reduce the number of animals needed to precisely determine mouse prostate volume changes.**

What were the acquisition times of the reported sequences? Each sequence seems to be rather long. I calculated about 150 minutes. This would mean that only one of these sequences could have been acquired per animal. Is that correct?

Yes, this is correct, image acquisition time was 2.5 hours. A single image was acquired at each timepoint.

Why have the total pixels to be divided by the magnification factor in order to generate the prostate volume? Please clarify.

We have removed any reference to image magnification from the manuscript.

Voxel size does not have to be explained extensively.

The first two sentences of the “Regression and re-growth…” section in results should be moved to the introduction or discussion.

These have been rewritten and moved.

The use of high field magnets and small coils to improve resolution is not new. Please rephrase in the discussion.

As far as we have been able to determine, it is novel in the mouse prostate, based on the literature available.

Please rephrase the statement in the discussion that the used MRI protocol accurately determined the volume of the prostate. This was not studied in this manuscript.

The term accurate has been removed.

Figures 1, 2, 4 and 5: One row of images is enough. The ones that outline the prostate can be interchanged with the ones that don’t outline it. Figure 1: Please mark some anatomic landmarks. What is the round structure in the middle of the gland?
Figure 1: Where is the bladder? Shouldn't there be a cross section of it?
Figure 1: Why is the prostate shaped so oddly? I would expect it to be more round as it is in humans and other animals.
Figure 4: Panels A to D can be removed. They do not provide any additional information.

Since the lines outlining the lobes of the prostate interfere with visualization of the margins, both the original images and the segmented images are presented for the reader. The identification of anatomical structures is quite clear in Figure 3 (the round structure is the ureter). In the mouse, the bladder is caudal to the prostate and appears on slices of the image stack not presented, as there is little or no prostate in those slices, and each image is the slice with the largest segmented area. The prostate is shaped oddly, as is described in the discussion and references therein, because it is interdigitated with surrounding organs in the mouse, unlike other species.

Discretionary Revisions

The last sentence of the section animals in the methods should be moved to the front.