Reviewer's report

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

Version: 1 Date: 15 February 2007

Reviewer: John Waterton

Reviewer's report:

General

General questions in assessing the work for BMC Urology:
1. Is the question posed by the authors new and well defined?
   o The question posed by the authors appears to be: "Considering the prostate volume in aged mice, as
   measured by an optimised MRI procedure, what is the test-retest reproducibility of the measurement, and
   what is the magnitude of the response to orchidectomy and DHT?".
   • The question has been addressed by previous authors (some cited) although not apparently in aged mice;
     previous studies have achieved similar resolution in mouse prostate.
   • The aged mouse model does not appear previously to have been employed with MRI volumetry
2. Are the methods appropriate and well described, and are sufficient details provided to replicate the work?
   o The orchidectomy and DHT substudies lack a sham control and seem underpowered
   o The methods are incompletely described
   o The statistical analysis of reproducibility is inadequate
3. Are the data sound and well controlled?
   o The orchidectomy and DHT substudies lack a sham control and seem underpowered
4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
   o The authors are invited to review the FRAME recommendations for reporting studies in living animals, and
     to consider whether adopting any of those recommendations would benefit the manuscript
5. Are the discussion and conclusions well balanced and adequately supported by the data?
   o The statistical analysis of reproducibility is inadequate
   o The orchidectomy and DHT substudies lack a sham control and seem underpowered
   o The use of the term "accuracy" is inappropriate as there is no comparison with a "gold standard" such as
     pathology.
6. Do the title and abstract accurately convey what has been found?
   o The title and abstract do not make it clear that this is a small incompletely controlled study
   o The abstract would benefit from specific statement of the reproducibility found
7. Is the writing acceptable?
   o The manuscript is clearly written in good scientific English with occasional errors in spelling and grammar
   o Some of the material is inappropriately allocated between "methods", "results", and "discussion"

The authors are to be commended on their mention of reduced animal use in the abstract as this will assist
researchers with an interest in animal welfare who are searching the literature for techniques to reduce
animal use

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

Major Compulsory Revisions (that the author must respond to before a decision on publication can be
reached)

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

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?

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?

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.

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, \( ?s \), for the repeat measurements on that subject, divided by the mean volume, \( ?s \), for the subject. Even when as few as two or three subjects is used then \( ?s \) 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

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?

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

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.

-------------------------------------------------------------------
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
Abstract implies that all volumetry used CHESS
Actual reproducibility should be stated – do not use "very"
Accuracy was not measured – use "precise"
"up to 10 imaging sessions over 5 weeks" – this is not clearly described in Methods
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.

Some of the text in the second paragraph belongs in "Methods"

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

There are no striped bars in my Fig 3

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

-------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)
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:
- in genetically modified mice (e.g. Degrassi 2007; Fricke 2006; Garcia 2006; Shukla 2005; Gupta 2004/2001; Abdulkadir 2001; Eng 1999; Hsu 1998)
- 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
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.

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.

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.

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 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).

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

Please discuss as far as possible any differences in volume with supplier, age, or breeder status.