Author's response to reviews

**Title:** Quantification of Bone Marrow Lesion Volume and Volume Change Using Semi-automated Segmentation: Data from the Osteoarthritis Initiative

**Authors:**

Jincheng Pang (pangjc@eecs.tufts.edu)
Jeffrey B Driban (jdriban@tuftsmedicalcenter.org)
Geoffroy Destenaves (GDestenaves@tuftsmedicalcenter.org)
Eric Miller (elmiller@ece.tufts.edu)
Grace H Lo (ghlo@bcm.edu)
Robert J Ward (rward@tuftsmedicalcenter.org)
Lori Lyn Price (lprice1@tuftsmedicalcenter.org)
John A Lynch (jlynch@psg.ucsf.edu)
Charles B Eaton (charles_eaton@mhri.org)
Felix Eckstein (felix.eckstein@pmu.ac.at)
Timothy E McAlindon (tmcalindon@tuftsmedicalcenter.org)

**Version:** 4  **Date:** 11 December 2012

**Author's response to reviews:** see over
Enclosed is a revised copy of a full-length article manuscript for *BMC Musculoskeletal Disorders* entitled "Quantification of Bone Marrow Lesion Volume and Volume Change Using Semi-automated Segmentation: Data from the Osteoarthritis Initiative." This manuscript was prepared by Jincheng Pang, Geoffroy Destenaves, Eric Miller, Grace Lo, Robert Ward, Lori Lyn Price, John Lynch, Charles Eaton, Felix Eckstein, Timothy McAlindon, and myself. All of the authors meet the Uniform requirements for Manuscripts Submitted to Biomedical Journal criteria for authorship. This manuscript has not been submitted and will not be simultaneously submitted to another journal. This study was supported by the NIH/NIAMS (grant 1R01AR054938).

This manuscript describes the construct validity of a novel semi-automated method of quantifying the volume of bone marrow lesions (BMLs), which are common findings in osteoarthritis. These lesions have been previously related to cartilage loss and joint symptoms. Unfortunately, previous methods have been time consuming which limits researchers’ ability to measure a large number of knees. This new method can reliably quantify BMLs in a set of knee magnetic resonance images with only 4 to 12 minutes of an investigator’s time. This BML segmentation approach may be an excellent opportunity to quantify BMLs in large data sets (e.g., the Osteoarthritis Initiative).

We have revised the manuscript based on the reviewers’ comments. Therefore, we are submitting a revised manuscript and a response to the reviewers’ comments that describes how the manuscript was updated (attached on subsequent pages).

I will be serving as the corresponding author on this study. My contact information is below:

Corresponding Author: Jeffrey Driban, PhD, ATC, CSCS  
Division of Rheumatology  
Tufts Medical Center  
800 Washington Street, Box #406  
Boston, MA 02111  
Email: jdriban@tuftsmedicalcenter.org  
Phone: 617-636-7449  
Fax: 617-636-1542

Thank you in advance for your time and commitment in reviewing this manuscript.

Sincerely,

Jeffrey B. Driban, PhD, ATC, CSCS  
Division of Rheumatology  
Tufts Medical Center
We would like to thank the reviewers for offering their time and feedback on the manuscript. We have modified the manuscript based on the reviewers’ comments and offered a response below in bold.

Reviewer: Dawn Doré
Reviewer’s report:

This study aimed to validate a new, three-dimensional, semi-automated BML segmentation method. Authors demonstrated that their volumetric BML measurements increased with a well-accepted semi-quantitative BML scoring system (BLOKS). They also showed that change in their BML volume measurements were associated with longitudinal cartilage loss over 24 months. Their method is time-efficient and demonstrated adequate reliability. Time-efficient BML measurements are very important, given the large number of participants in OA cohort studies. This paper would be of great interest to the OA research community.

I have a few minor comments/questions:

1) Can you expand on the rationale for omitting the middle 9 slices? Wouldn’t you get BMLs adjacent to cartilage surface on these slices as well?
   We omitted the middle 9 slices because of difficulties getting accurate segmentation results in this region because of challenges segmenting the border of the bone and increased signal-intensity heterogeneity. These images correspond to the tibial spines but you are correct that in the femoral region the BMLs towards the anterior femur would be adjacent to cartilage. The decision to exclude these slices introduces a bias in which patellofemoral BMLs may be under represented but provides a good measure of BMLs that may be influencing the tibiofemoral joint. Hernandez-Molina et al (2008) provide evidence that centrally located BMLs are not associated with changes in tibiofemoral cartilage unless they extend into the medial compartment (which this segmentation approach would detect).
   The rationale has been clarified in the manuscript (Methods, Semi-automated BML Segmentation, 2nd paragraph) and the discussion has been modified to further clarify the potential limitation (Discussion, 7th paragraph).

2) How does the program distinguish between a BML and a cyst?
   This program does not differentiate regions of BMLs and cysts. We acknowledge that this may be a limitation in the discussion (7th paragraph).

3) How does the slice thickness of 3 mm impact on the volume measurements?
   We hypothesize that a slice thickness of 3 mm may be less sensitive to detect changes than smaller slice thicknesses; however, this has not been tested yet. Despite the 3 mm slice thickness we demonstrated good construct validity.

4) Why haven’t the authors adjusted for any covariates in their analysis between BMLs and cartilage damage? Age, sex, BMI, radiographic OA status? This may be due to a small sample size, please explain.
   We did not adjust these analyses since the goal was to replicate previously reported associations between two measurements to determine the construct validity of the new BML segmentation method. In this situation, we were not concerned about potential covariates. This has been clarified in the Methods section (Study 2, 2nd paragraph).
5) I assume there were a lot of zero’s for the BML volume measurements, i.e. those who did not have a BML at a specific site. In the analysis with longitudinal cartilage loss, how did the authors deal with this?

   In these analyses every knee had a measured non-zero baseline BML volume in the tibia (range = 18.0 mm$^3$ to 8120.3 mm$^3$) and femur (range = 21.8 mm$^3$ to 4205.4 mm$^3$). This data has been added to Table 1.

6) Both validation studies in this paper were done on participants with denuded cartilage/full thickness cartilage loss. This could have strengthened the association between BML volume change and longitudinal cartilage loss. It is reassuring that this method demonstrated validity in such small numbers (n=48 and n=38); however, it does need further exploration in a larger cohort, with varying degrees of cartilage damage (i.e in those with healthy knee’s as well).

   We agree that the inclusion of only knees with denuded area strengthened the association which was ideal in initial validation studies. In the conclusion we added that the association was limited to knees with full thickness cartilage loss. Additionally, the final paragraph of the discussion has been updated to note that the validity of this method may need to be re-evaluated when deployed in different study populations (e.g., knees without full thickness cartilage loss).

7) This study does validate a semi-automated method to measure BML volume. BML volume may be a more accurate method to measure BML size; however, this remains to be determined. It would be interesting in future studies to see whether BML volume is more strongly associated with factors such as cartilage loss, pain, joint replacement, compared to a semi-quantitative or quantitative (2D)measure of BML size. If so, then it would be the optimal BML measure to use. Overall this is an interesting paper. BMLs are recognised as a key feature in OA and to have a program which can accurately measure their size and size changes in a time-efficient manner is invaluable.

   Thank you.

Reviewer: Daichi Hayashi
Reviewer’s report:

Major Compulsory Revisions:
1. In the Background section, authors cited a study by Tanamas et al (ref 2, Rheumatology Oxford 2010), but this study has a significant shortcoming in the methodology regarding BML assessment using MRI. Basically, the authors used T1-weighted fat-suppressed 3D gradient echo-type sequence, which is known to be inadequate. Correct pulse sequences should always be used for reliable BML assessment on MRI. Thus, the reference 2 should be deleted. Instead, authors can cite a much more important publication describing the relationship between BML and pain in OA and their fluctuation: Zhang Y et al. Arthritis Rheum 2011;63:691-9.
2. The same methodological flaw is evident in reference 5 (Raynauld JP et al. ARD 2008). This citation should also be deleted from the manuscript. The authors of the submitted manuscript did use the correct pulse sequence for the present study, so it is surprising to see that they are citing two studies that have significant methodological flaws.
3. In Discussion, page 12, second paragraph, authors cite reference 5 again to support their discussion. However, as I noted earlier, this reference should not be cited because it will not support the authors’ claim in a scientifically meaningful manner. Please cite alternative publication instead of ref. 5.

   Points 1-3: We have added the reference for Zhang Y et al. Arthritis Rheum 2011;63:691-9. We appreciate the reviewer’s concerns and acknowledge that the sequence the authors of these references used is less sensitive and systematically under
estimates BML size compared to the standard sequence. Despite this limitation, these authors have demonstrated reliability with their measures and have generated similar results to prior work with optimal sequences, suggesting that despite a systematic bias their results may still be valid. We have opted to keep these references because their findings along with other papers support the statement we are referencing; however, if the editor wishes we'll remove these references.

Minor Essential Revisions:
Abstract:
4. In the Methods section, please state what statistical method was used to derived z and p values described in the Results section. Was p<0.05 considered statistically significant?

The methods section of the abstract and manuscript have been clarified (Study 1, 2nd paragraph; Study 2, 2nd paragraph).

5. Please concisely state in the Methods section how the proposed 'new' segmentation method is different from pre-existing one.

We have updated the discussion (1st paragraph) to briefly describe how the new segmentation method is different from pre-existing methods.

6. In the results section, please state how many BMLs were identified/involved in each study.

In study 1, 2 knees had tibial BML volume equal to zero. This has been added to the results section (1st paragraph). In study 2, every knee had tibial BML volumes.

Background:
7. While the reference 7 (Peterfy CG, et al. OAC 2004) does indeed mention the appropriate choice of pulse sequences for semiquantitative assessment of various OA features, it is not specific to bone marrow lesions. Perhaps it would be better that the citation of ref 7 (page 3, last line) be replaced by a more specifically BML-oriented imaging paper, such as: Xu L, et al. Semin Arthritis Rheum 2012;42:105-18.

We agree that Peterfy CG, et al. OAC 2004 does not mention a specific sequence for bone marrow lesions but this paper is referenced since the authors state “subarticular bone marrow abnormality was defined as poorly marginated areas of increased signal intensity in the normally fatty epiphyseal marrow on fat-suppressed T2-weighted FSA images.” We also included the reference to your paper Xu L et al 2012.

Methods:
8. In page 8, the first paragraph of "Study 1" section, authors state: "We selected 80 right knees with ........and had acceptable quality fixed-flexion knee radiograph and MR imaging sequences (identified as read project 4 in kmri_qcart_ecksteinXX [version 0.4 and 3.3])." This last part, "identified as read project 4 in kmri_qcart_ecksteinXX" does not mean anything to readers who are unfamiliar with specific terminology used by Dr. Eckstein's research group. Please cite appropriate reference(s) so that it can be understood by general readers who are (mostly) not an expert of quantitative cartilage morphometry.

It has been clarified that this information is needed to properly select the data from the public online files (Study 1, 1st paragraph).

9. Please explain why authors selected to use BLOKS as the semiquantitative scoring method of their choice. Why not other scoring systems that are available in the literature?

We have added an explanation for why we selected the BLOKS scoring method (Study 1, 1st paragraph). “We selected BLOKS scores, instead of other BML scoring methods, based on our experience with BLOKS scoring and prior research which
indicated that BLOKS was comparable with Whole Organ MR Scoring (WORMS) method in cross-sectional studies (Lynch JA et al., 2010 OAC).

10. The selection criteria/exclusion criteria should be concisely presented in a flow-chart, which would greatly help readers of this paper. Please use a flow chart to explain how the authors narrowed down their sample from a vast collection of OAI database.

   In the text we have clarified that these studies used two convenience samples (Study 1, 1st paragraph; Study 2, 1st paragraph). While we agree that flow charts would be a good complement to the selection criteria described in the text we are concerned that adding two more figures to the manuscript may take a lot of space and distract from the purpose of this manuscript.

Results:
11. I suggest authors include a table which summarizes the results of pertinent statistical analyses. It is easier to see them in a table, rather than reading through the text only.

   We agree that some of the results section could be better presented in a table; therefore, we added Table 1.

Conclusion:
12. Authors state that "this new method will enable researchers to assess larger MR data sets in a time efficient and cost effective manner." However, authors did not present any data regarding the cost in this paper. Thus, this statement is not supported by the evidence and should be modified accordingly.

   We agree. The sentence has been modified to state that "this new method will enable researchers to assess larger MR data sets in a time efficient manner" (Conclusion Paragraph).

Figures:
13. Figure 1: Each figure parts should be cropped so that the region of interest can be magnified. Authors should crop the upper and lower parts (proximal femur and distal tibia), as well as left and right margins of the images (black space and fat in front of patella, and fat behind the posterior compartment).

   We agree and have reformatted Figure 1.

14. Figure 2: Within this figure, authors wrote "Distribution of v0vol by tib". This is a specific terminology for their analysis, and is not universally understood by general readers. Please replace this notion by a more generally understandable term.

   We agree and Figure 2 has been updated by omitting the title of the figure since the Figure legend provides this in a more effective manner.