Ms. Ref. No.: AANAT2032
Title: INTRA- AND INTER-OBSERVER RELIABILITY OF QUANTITATIVE ANALYSIS OF THE INFRA-PATELLAR FAT PAD AND COMPARISON BETWEEN FAT- AND NON-FAT-SUPPRESSED IMAGING - DATA FROM THE OSTEOARTHRITIS INITIATIVE
Annals of Anatomy

Received	Aug 12, 2015
Decision	Aug 19, 2015
Revision received	Oct 05, 2015
Accepted	Oct 09, 2015

Decision letter

Dear Dr. Steidle-Kloc,

The reviewers have commented on your above paper. They indicated that it is not acceptable for publication in its present form.

However, if you feel that you can suitably address the reviewers' comments (included below), I invite you to revise and resubmit your manuscript.

The comments provided by the reviewers very often considerably help to improve and strengthen a paper. In addition, the authors' response letter accompanying a revised version often contains important information which may be regarded as an added value to the final version of the manuscript. Therefore, Editors decided that the non-confidential comments of the reviewers and the non-confidential authors' responses will be published as online supplementary material together with the final version of an eventually accepted article.

Please carefully address the issues raised in the comments.

If you are submitting a revised manuscript, please also:

a) outline each change made (point by point) as raised in the reviewer comments

AND/OR

b) provide a suitable rebuttal to each reviewer comment not addressed

I look forward to receiving your revised manuscript.

Yours sincerely,

Friedrich Paulsen
Editor-in-Chief
Annals of Anatomy

Reviewers' comments:

Reviewer #1: INTRA- AND INTER-OBSERVER RELIABILITY OF QUANTITATIVE ANALYSIS OF THE INFRA-PATELLAR FAT PAD (IPFP) AND COMPARISON BETWEEN FAT- AND NON-FAT-SUPPRESSED IMAGING - DATA FROM THE OSTEOARTHRITIS INITIATIVE

General Comments:

This study aimed at analyzing inter- and intra-observer reliability of Hoffa fat pad segmentations from intermediate-weighted fat suppressed and T2-weighted non fat suppressed MR images acquired in the OAI cohort. As Hoffa's fat pad has come increasingly into the focus of the osteoarthritis research community the study is of high relevance. The methodology applied is adequate to pursue these aims. One of the few remaining questions is the intra-reader reliability being performed on studies acquired one year apart. Clinical experience and large on-going studies show marked fluctuation in Hoffa fat pad size commonly due to different amounts of concomitant joint effusion albeit not much literature is available on the topic. The authors are encouraged to discuss this in more detail.

In addition, the issue of QC/over-reading of segmentations needs to be discussed in more detail, i.e. whether better reliability may be expected using experienced radiologists or imaging experts rather than junior researchers or whether the authors believe that QC is paramount regardless of the experience of the initial readers.

Finally the low n of 9 should be discussed in more detail.

Specific Comments:

Introduction

Page 4 l 19: "Even after 20 normalizing IPFP volume to body weight, men still displayed a 9% greater ratio" Please explain what ratio is meant.

First part is relatively lengthy given the somewhat controversial data available on the topic. The second part that leads to the research question of a comparison of non-fs vs. fs images is relatively short and should be supported by references, e.g. why non-fs imaging should be more suitable for segmentation than fs imaging. Please mention in the introduction if you are referring to T1- or T2-weighted imaging.

Methods:

P6 line 53: Please change "incident cohort" to "incidence cohort"

P6 l75: Please briefly give reason that for comparability with the IW-fs sequencethe acquisition with 30 ms TE from the MESE was chosen for the segmentations.

P7 l 88/89: please describe what is meant by "label" in this context.

L96-100: Motivation to use baseline and 12 months images to test intra-observer realibilty is unclear. Why were not the same baseline images re-segmented? There is some definite variation in Hoffa fat pad morphology over one year especially depending on the amount of associated intraartcular joint fluid (e.g. recesses).

Discussion

L181: Please delete "as" in first sentence of discussion.

L184: please add "fluid-sensitive" before "fs".

L185 please add "by junior researchers" after the word "imaging". The differences in segmentations of experienced e.g. radiologists or other imaging experts and junior researchers has not been tested. It would be interesting to understand if expert QC is a prerequisite for adequate reliability regardless of imaging experience of the segmenters or not. Please comment.

Figure 1 legend: please add "intermediate-weighted" in front of "fat suppressed".

Tables 2 and 3: please discuss briefly whether all 6 measures of IPFP quantitative characteristics are independent measurements.

Reviewer #2: Dear authors

Congratulations to this remarkable manuscript. It was a pleasure to read.
The manuscript as a whole is in line with the "state of the art" research and leaves no doubt about the hard work and good understanding of the field of anatomy and imaging of the authors.
Following, I would suggest slight changes, which have to be clearly seen as suggestions.
Heading
Maybe delete the abbreviation. It is not really needed here.
Introduction
Line 34: maybe change "these" for "they"?
Line 36: maybe put "therefore" in the beginning of the sentence.
Materials and methods
Line 48: here is the abbreviation for MRI, but the "MRI" was used in line 32, 38 before. The first use was in line 18 of the introduction, so maybe you will abbreviate it there.
Line 63: Please change (4 men, 5 women) to (5 women, 4 men) as it would be in order with your f/m in your table and this way is easier to associate.
Line 74: please change Fig. 1 A-B to Fig. 1 A,B since there is nothing in between A and B
Line 77: the same
Line 82, 83: E.S. Furthermore, "who were a postdoc" that does not sound right, please rephrase and put (J.D. ….) directly behind the students and not behind the University.
Line 88 and 89: "the one" is not necessary, maybe delete?
Results
Line 123: Why is Table 1 in bold font?

The rest: (Table 2) is quite often in your section. Firstly, tables define themselves as giving additional data and not repeating the results-section, but still I would not change anything here, just to mention. But it might be a good idea to maybe get rid of some (Table 2) since they are sometimes even at the end of following sentences.
Discussion
Line 244: synovial tissues, such as meniscus. First, I would suggest to use menisci, the plural. Second, menisci are no synovial tissue. Please rephrase.
Line 248: IPFP imaging biomarkers may… That just makes no sense. IPFP imaging as biomarker… maybe…
Line 250 and 252 are the same references. Please only quote once.
Conclusion
Line 6: makes no sense since you still have the "compared" in the sentence which probably was your first idea. Use the same as in the summary.
Abbreviation
I know it sounds silly to do, but MRI, OA and 3D have to be put in there.
Figures and Tables
Fig. 2 B
The pictures are good but the writing and the arrow is destroying your picture (the arrow is not straight). Why not put "left=anterior" in the legend and get rid of "anterior surface" and just put an orange line onto the blue volume and state in the legend: "blue=volume", "orange=depth"?
Fig. 3
Each graph shows 3xcm. I think I understood that you are measuring in cm. One would be more than enough.
Since your x and y axes are all labelled the same, I would suggest to define them in the legend and get rid of all the writing in the graphs, just leaving the heading A) IPFP volume… and so on.
You could even argue, since IPFP is stated in the legends if it is needed in the graphs at all, leaving you just with A) Volume (cm3)
That would look much nicer.
Furthermore, try to make the "A" in Fig 3 the same size as the "A" in Fig 2 and please use the same font which you did not.

Table 1
You state "Follow- up (Y1)" (column 3) and afterwards use Y1 FU (column 4-6) as abbreviation. That is confusion. Please just use Y1 since that is enough.
Furthermore you describe the abbreviations below the table but don't say what BL is. It seems logic at the second glance, but please add that either in the abb. below or in the heading of the first column.
Your abbreviations section at the bottom stops with a ;
Table 3
Here, you abbreviate FU and BL but don't use it.

Authors response letter

Reviewer #1:
General Comments:

This study aimed at analyzing inter- and intra-observer reliability of Hoffa fat pad segmentations from intermediate-weighted fat suppressed and T2-weighted non-fat-suppressed MR images acquired in the OAI cohort. As Hoffa's fat pad has come increasingly into the focus of the osteoarthritis research community the study is of high relevance. The methodology applied is adequate to pursue these aims.

Author comment: Thank you very much for the positive feedback.

One of the few remaining questions is the intra-reader reliability being performed on studies acquired one year apart. Clinical experience and large on-going studies show marked fluctuation in Hoffa fat pad size commonly due to different amounts of concomitant joint effusion albeit not much literature is available on the topic. The authors are encouraged to discuss this in more detail.

Author comment: Thank you for the thoughtful comment.

Author action: The following text was added to page 11, line 211-215.
In longitudinal studies in OA patients, IPFP size may be potentially influenced by variability in concomitant effusion. Future longitudinal studies on the effect of effusion on IPFP volume should thus take into account the normal variability observed here in healthy OAI participants over a one year period, in which effusion was unlikely due to the lack of risk factors of OA.

In addition, the issue of QC/over-reading of segmentations needs to be discussed in more detail, i.e. whether better reliability may be expected using experienced radiologists or imaging experts rather than junior researchers or whether the authors believe that QC is paramount regardless of the experience of the initial readers.

Author action: To discuss the issue of QC/over-reading of segmentations in more detail the following text was added to page 13, line 252-257.
Relatively inexperienced (junior) researchers participated in the study, as these are commonly involved in studies requiring large scale segmentation. It is well possible that the inter-observer reliably before QC would have been higher, if experienced radiologists or more experienced imaging experts had been involved. Whether or not QC is paramount regardless of the experience of the readers will have to be shown in separate studies, but for the time being, a rule-based approach ….

Finally the low n of 9 should be discussed in more detail.

Author action: We have added the following text on page 11, line 217-220 to explain in detail the n that was used in this study.
Although a limitation of this study is the relatively small sample size of nine data sets, the degrees of freedom of the present inter-observer analysis is 9 x (5-1) = 36, which is above the minimum of 27 degrees of freedom recommended by Glüer et al. (Glüer

et al., 1995), insuring that the real precision error is not underestimated by more than 30%.

Specific Comments:

Introduction

Page 4 l 19: "Even after normalizing IPFP volume to body weight, men still displayed a 9% greater ratio" Please explain what ratio is meant.

Author action: We have modified the text on page 3, line 22 to explain the ratio more clearly.
Even after normalizing IPFP volume to body weight, men still displayed a 9% greater ratio of IPFP volume/body weight than women.

First part is relatively lengthy given the somewhat controversial data available on the topic. The second part that leads to the research question of a comparison of non-fs vs. fs images is relatively short and should be supported by references, e.g. why non-fs imaging should be more suitable for segmentation than fs imaging.

Author comment and action: The first part of the introduction on the topic was deliberately kept on the longer side, as readers of Annals of Anatomy may not be familiar with that topic studied here. The purpose of the introduction therefore was to provide at least a short review of the research context, in which this methodological study is set.
We have now extended the second part of the introduction to provide a better rationale of the research question of a comparison of nfs vs. fs images. We are not aware of literature making a specific recommendation on which MRI sequence should be used for the purpose of IPFP quantification, but intuitively, non-fat-suppressed (nfs) MRIs appear most suitable for this purpose, because in these adipose tissues are displayed with high signal intensity and good contrast to neighboring non-adipose tissues, such as bone cortices, ligaments, menisci, and cartilage. We now stated on page 4, line 34-41.
Intuitively, non-fat-suppressed (nfs) MRIs appear most suitable for the purpose of measuring the IPFP, because in these adipose tissues are displayed with high signal intensity and good contrast to neighboring non-adipose tissues, such as bone cortices, ligaments, menisci, and cartilage. However, many epidemiological studies preferentially acquire fat-suppressed (fs) knee MRIs, because they are better suited for evaluating structural pathology of articular tissues, such as the meniscus, cartilage and bone marrow (Guermazi et al., 2013), and it has not been previously shown whether these permit to analyze the IPFP with similar reliability as non-fs images.

Please mention in the introduction if you are referring to T1- or T2-weighted imaging.

Author comment and action: In the introduction, we neither refer to T1 nor to T2 weighted images, but to intermediate-weighted MRIs. We have added this in the text on page 4, line 43.
…and to compare quantitative measures of the IPFP as obtained from intermediate-weighted fs versus nfs MRIs.

Methods:
P6 line 53: Please change "incident cohort" to "incidence cohort"

Author action: The word “incident” was replaced by “incidence”.

P6 l75: Please briefly give reason that for comparability with the IW-fs sequence the acquisition with 30 ms TE from the MESE was chosen for the segmentations.

Author action: We have now added an explanation to the text on page 6, line 85-87.
The 30 ms time of echo acquisition of the MESE was selected in order to provide consistency in that parameter between the fs and nfs acquisition.

P7 l 88/89: please describe what is meant by "label" in this context.

Author action: We have tried to clarify the text on page 6, line 98-99 to state.
The anterior border of the IPFP (facing the lig. patellae) was segmented using one label (green marker) and the posterior border (the one facing the knee joint) using another label (magenta marker; Fig. 2).

L96-100: Motivation to use baseline and 12 months images to test intra-observer reliability is unclear. Why were not the same baseline images re-segmented? There is some definite variation in Hoffa fat pad morphology over one year especially depending on the amount of associated intra-articular joint fluid (e.g. recesses).

Author action: We have now clarified the rational for analyzing images one year apart by adding text to page 11, line 211-215.
In longitudinal studies in OA patients, IPFP size may be potentially influenced by variability in concomitant effusion. Future longitudinal studies on the effect of effusion on IPFP volume should thus take into account the normal variability observed here in healthy OAI participants over a one year period, in which effusion was unlikely due to the lack of risk factors of OA.

Discussion
L181: Please delete "as" in first sentence of discussion.

Author action: The word “as” was removed.

L184: please add "fluid-sensitive" before "fs".

Author comment: We apologize for being unclear. The abbreviation “fs” stands for fat-suppressed (fs) images and not for ”fluid-sensitive”.

L185 please add "by junior researchers" after the word "imaging".

Author action: As recommended by the reviewer the text was added.
Therefore, the current study was performed to evaluate the inter- and intra-observer reliability of IPFP morphology, and to compare IPFP measures derived from more commonly used fs than nfs MR images by relatively inexperienced (junior)

researchers, who are commonly involved in studies requiring large scale segmentation.

The differences in segmentations of experienced e.g. radiologists or other imaging experts and junior researchers has not been tested. It would be interesting to understand if expert QC is a prerequisite for adequate reliability regardless of imaging experience of the segmenters or not. Please comment.

Author action: Thank you for the thoughtful comment. The following text was added to page 13, line 252-257.
Relatively inexperienced (junior) researchers participated in the study, as these are commonly involved in studies requiring large scale segmentation. It is well possible that the inter-observer reliably before QC would have been higher, if experienced radiologists or more experienced imaging experts had been involved. Whether or not QC is paramount regardless of the experience of the readers will have to be shown in separate studies, but for the time being a …

Figure 1 legend: please add "intermediate-weighted" in front of "fat suppressed".

Author action: We have modified the legend of figure 1.
Intermediate-weighted fat-suppressed and non-fat-suppressed sagittal MR images of the infra-patellar fat pad (IPFP).

Tables 2 and 3: please discuss briefly whether all 6 measures of IPFP quantitative characteristics are independent measurements.

Author action: We have now added some text in the discussion on page 12, line 237-240 to address this point.
While the different parameters computed here are not independent and correlated to each other, future studies will have to show which of these have the highest sensitivity to change in longitudinal trials, and which approach is thus the most efficient in large-scale applications.

Reviewer #2: Dear authors

Congratulations to this remarkable manuscript. It was a pleasure to read.
The manuscript as a whole is in line with the "state of the art" research and leaves no doubt about the hard work and good understanding of the field of anatomy and imaging of the authors.

Author comment: Thank you very much for the favorable evaluation.

Following, I would suggest slight changes, which have to be clearly seen as suggestions.
Heading
Maybe delete the abbreviation. It is not really needed here.

Author comment: We are very happy to follow these suggestions. The abbreviation in the heading has been removed.
INTRA- AND INTER-OBSERVER RELIABILITY OF QUANTITATIVE ANALYSIS OF THE INFRA-PATELLAR FAT PAD AND COMPARISON BETWEEN FAT- AND NON-FAT-SUPPRESSED IMAGING – DATA FROM THE OSTEOARTHRITIS INITIATIVE

Introduction
Line 34: maybe change "these" for "they"?

Author action: The word “these” was replaced by “they” on page 4, line 38 (line numbers changed because of the added text).

Line 36: maybe put "therefore" in the beginning of the sentence.

Author action: The word “therefore” was added on page 4, line 41.

Materials and methods
Line 48: here is the abbreviation for MRI, but the "MRI" was used in line 32, 38 before. The first use was in line 18 of the introduction, so maybe you will abbreviate it there.

Author comment and action: We apologize for the oversight and have changed the text on page 5, line 54 as recommended.

Line 63: Please change (4 men, 5 women) to (5 women, 4 men) as it would be in order with your f/m in your table and this way is easier to associate.

Author comment and action: Thank you for spotting this inconsistency. The text on page 5, line 69 has been changed.
The reliability of IPFP morphology in this study was performed in nine right knees (5 women, 4 men) of the OAI healthy reference cohort.

Line 74: please change Fig. 1 A-B to Fig. 1 A,B since there is nothing in between A and B

Author action: Changes have been made on page 6, line 80, 83 as recommended.

Line 77: the same Line 82, 83: E.S. Furthermore, "who were a postdoc" that does not sound right, please rephrase and put (J.D. ….) directly behind the students and not behind the University.

Author action: The text was rephrased on page 6, line 90-92.

Manual segmentation of the knees was performed by one postdoc (E.S) and four medical students (J.D., A.P., T.P. and F.H.) of the Paracelsus Medical University.

Line 88 and 89: "the one" is not necessary, maybe delete?

Author action: Changes have been made on page 6, line 97 as recommended by the reviewer.
The anterior border of the IPFP (facing the lig. patellae) was segmented using one label….

Results
Line 123: Why is Table 1 in bold font?

Author action: We apologize for the oversight and have changed the style on page 8, line 132.
The demographic data of the nine OAI healthy reference subjects are shown in Table 1.

The rest: (Table 2) is quite often in your section. Firstly, tables define themselves as giving additional data and not repeating the results-section, but still I would not change anything here, just to mention. But it might be a good idea to maybe get rid of some (Table 2) since they are sometimes even at the end of following sentences.

Author comment and action: We agree with the reviewer and followed his suggestion. The repeated "(table 2)" were removed.

Discussion
Line 244: synovial tissues, such as meniscus. First, I would suggest to use menisci, the plural. Second, menisci are no synovial tissue. Please rephrase.

Author action: As recommended by the reviewer the text has been rephrased on page 13, line 270.
This is important, because the MR imaging protocols of many epidemiological studies are restricted to fs imaging, as these images are generally better suited for evaluating structural pathology of synovial joint tissues, such as menisci, cartilage or bone marrow.

Line 248: IPFP imaging biomarkers may… That just makes no sense. IPFP imaging as biomarker… maybe…

Author comment and action: Thank you for your suggestion. We have modified and rephrased the test on page 13, line 274.
IPFP imaging measures as a potential biomarker may be of particular interest in clinical trials that explore non-pharmacological interventions in knee OA.

Line 250 and 252 are the same references. Please only quote once.

Author action: We apologize for repeating the reference and have removed it as recommended on page 13, line 276.

It has recently been reported that IPFP volume may be responsive to exercise and/or diet as treatment of knee OA, with the combination of exercise and diet being most effective in reducing IPFP volume (Pogacnik Murillo et al., 2015).

Conclusion
Line 6: makes no sense since you still have the "compared" in the sentence which probably was your first idea. Use the same as in the summary.

Author comment and action: We apologize for being unclear. The conclusion on page 14 has now been adapted to that in the abstract.
The results suggest that quantitative measures of IPFP morphology can be performed with satisfactory reliability when expert QC is implemented. The IPFP is more clearly depicted in nfs images, and there is a small systematic off-set versus analysis from fs images. However, the high linear relationship between fs and nfs imaging suggests that fs images can be used to analyze IPFP morphology, when nfs images are not available.

Abbreviation
I know it sounds silly to do, but MRI, OA and 3D have to be put in there.

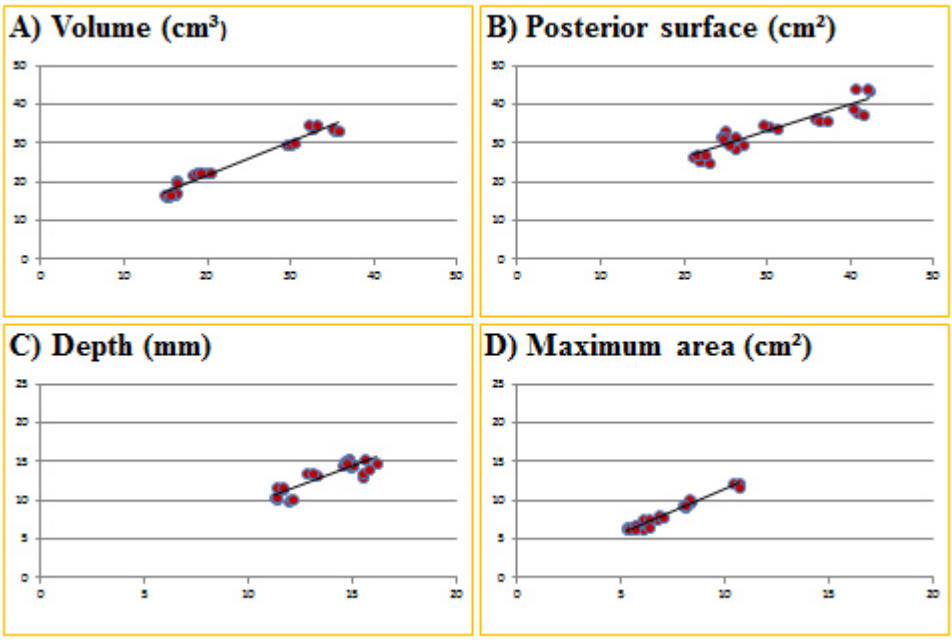
Author action: We apologize for the oversight. MRI, OA and 3D have been added to the list of abbreviations.

Figures and Tables
Fig. 2 B
The pictures are good but the writing and the arrow is destroying your picture (the arrow is not straight). Why not put "left=anterior" in the legend and get rid of "anterior surface" and just put an orange line onto the blue volume and state in the legend: "blue=volume", "orange=depth"?

Author comment and action: Thank you for your suggestion. The figure has been modified.
Legend: A) Sagittal fat-suppressed MRI of the knee joint showing segmentation of the anterior surface (green label), and posterior surface (magenta label) of the IPFP; B) IPFP volume (blue) and depth (orange line); C) 3D reconstruction of the IPFP viewed from posterior-lateral and D) from anterior-lateral;

Fig. 3
Each graph shows 3xcm. I think I understood that you are measuring in cm. One would be more than enough.
Since your x and y axes are all labelled the same, I would suggest to define them in the legend and get rid of all the writing in the graphs, just leaving the heading A) IPFP volume… and so on.
You could even argue, since IPFP is stated in the legends if it is needed in the graphs at all, leaving you just with A) Volume (cm3) That would look much nicer.
Furthermore, try to make the "A" in Fig 3 the same size as the "A" in Fig 2 and please use the same font which you did not.

Author action: The recommendations of the reviewer have been implemented in the figure.



Legend: Linear relationship between fat suppressed TSE (x-axis) and non-fat suppressed MESE (y-axis) sequence for the infra-patellar fat pad A) volume, B) posterior surface, C) depth and D) maximum sagittal area;

Table 1
You state "Follow- up (Y1)" (column 3) and afterwards use Y1 FU (column 4-6) as abbreviation. That is confusion. Please just use Y1 since that is enough.

Author comment and action: We apologize for being confusing. Table 1 was adapted to the reviewer comment: Follow-up was abbreviated as Y1.

Furthermore you describe the abbreviations below the table but don't say what BL is. It seems logic at the second glance, but please add that either in the abb. below or in the heading of the first column.

Author action: As recommended by the reviewer BL was added to the abbreviations.

Your abbreviations section at the bottom stops with a ;

Author action: The typo has been corrected.

Table 3 Here, you abbreviate FU and BL but don't use it.

Author action: As recommended by the reviewer BL and Y1 was added to the table and adapted in the abbreviations.