PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Clinical values of FDG PET in polymyositis and dermatomyositis syndromes: imaging of skeletal muscle inflammation
AUTHORS	Tateyama, Maki; Fujihara, Kazuo; Misu, Tatsuro; Arai, Akira; Kaneta, Tomohiro; Aoki, Masashi

VERSION 1 - REVIEW

REVIEWER	Kei Ikeda Chiba University Hospital, Japan
REVIEW RETURNED	07-Nov-2014

GENERAL COMMENTS	No further comments

REVIEWER	Takashi Kanda Department of Neurology and Clinical Neuroscience
REVIEW RETURNED	09-Nov-2014

GENERAL COMMENTS	Most of the points suggested by the referee in the previous version have been adequately responded. Although the problems according to the original study design are still present, most of them were described as the limitations of this study.

REVIEWER	David Hilton-Jones John Radcliffe Hospital Oxford, UK
REVIEW RETURNED	11-Nov-2014

GENERAL COMMENTS	Although the authors have provided responses to the many queries raised by the reviewers, a major consequence has been extension of comments about "limitations of the study". I have concerns about the relatively small numbers within each category of disease, the use of an antiquated classification system (Bohan & Peter) and lack of "useful" control subjects (i.e. muscular dystrophy, which is a major differential diagnosis). There is little evidence to suggest added-value of PET over existing methodologies.

VERSION 1 – AUTHOR RESPONSE

Response to the reviewers' comments

Reviewer 1

Major concerns

1. *According to Introduction, purpose of this study is to establish a simple and reliable way to assess muscle inflammation using FDG-PET. However, conclusions of Abstract and Discussion do not correspond to this purpose and only refer to the general utility of FDG- PET in PM/DM. Little is discussed about advantage and application of their own methods in the Discussion section either. Moreover, it is unclear whether Authors are proposing to use only vFDG-based methods or both vFDG-Based and SUV-based methods as a simple and reliable way.*
   We made changes in Abstract (page 2, line 4 and 22), Introduction (page 5, line 8-9, 14), Discussion (page 15 , line 15 - 18), and Strength and limitation of the study (page 3, line 4-5).

A. *If Authors are to demonstrate the utility of vFDG-based methods, using SUV-based methods as gold standard, the gold standard needs to be more comprehensive. In their study, SUV was not calculated in sternoclidomastoid muscles. Paraspinal muscles, or muscles in buttocks, where substantial proportions of patients had FDG uptakes (Table 1). SUV was not calculated in trapezius or deltoid muscle usually included in manual muscle tests (MMTs), a standard instrument in the assessment of myopathy. In fact, the lack of significant correlation of the mean SUVmax with the serum CK level, which is another standard instrument in the assessment of myopathy, may be due to the inappropriate selection of muscles.*
   We compared vFDG and SUV in our patients to evaluate an accuracy of the vFDG-based method. We found strong correlation between the mean SUVmax and the extent of vFDG positive regions. A comprehensive measurement of SUV was already demonstrated previously (Tanaka S et al. Rheumatology). They found a significant but moderate correlation between mean SUV and both MMT and CK. In our study, we measured SUV in proximal muscles of four extremities and found that the mean SUVmean of PM/DM in our study was comparable with that of Tanaka et al's. We think that comprehensive SUV calculation is not mandatory in our study.
   Serum CK level is a good marker of disease activity and useful to monitor the efficacy of therapies in each patient. However, an extent of CK elevation is variable among each patient and serum CK levels not always reflect disease severity. For example, some patients show preclinical hyper CKemia with little weakness and some patients show severe disease phenotype with only mild CK elevation. In this study, we did not find correlation between serum CK level and the total histological score ($p = 0.9296$). Instead, total histological score correlated with both the mean SUVmax ($p = 0.0002$) and the number of vFDG regions ($p = 0.0038$). We added sentences in Results (page10, line 21-23) and Discussion (page 13, line 1-4).

B. *If authors are proposing that both methods are useful, validated comparators are needed. Serum CK levels and histopathological scores in a single muscle do not sufficiently serve as the gold standard for global myositis activity. MMT score is a minimal requirement.*
   MMT was not scored in all the patients. A part of the patients were referred from other department for histological assessment. The patients in this study were able to do their daily life independently with some disability (page 6 , line 11-13).

2. *Anyways, the lack of data on muscle strength needs to be discussed as a major limitation of this study.*
   We described this point in the limitation of the study (page 3, line9-10, page 15, line 9-10).

3. *As for the SUV-base assessment, the data for both mean SUVmax and mean SUVmean are presented in an arbitrary manner. This is confusion and whether Authors find one measure more useful than the other or find both measures useful is unclear. The data for both measures should be presented in a more organized way so that their argument is clarified.*
   We assessed mainly SUVmax, because it may more contribute to vFDG. In the previous study by Tanaka et al, SUVmean was used for their evaluating method. For comparison our data with those of Tanaka et al. we also calculated SUVmean (page 12, line 8 - 10). We made changes in Results (page 10 , line11 - 12).

4. *Although vFDG-based assessment is interesting and can be useful. Authors only give information on the distribution of vFDG-positive muscle regions and only analyzed the number of vFDG-positive muscle regions. Given that this is a novel scoring system for PM/DM, the frequency of vFDG intakes equivalent to the mediastinum vessels (1 point) and that equivalent to the liber (2 points) should be provided separately. If the frequency of muscle which gained 2 points is very small, this methods can be even more simplified by focusing on the mediastinum vessels only.*
   According to the reviewer's comment, we simplified our methods; only mediastinum blood vessels as a positivity criterion. We made changes in Abstract (page 2, line 8), Methods (page 7, line 3-7). The results of statistical analysis were also changed in Abstract (page 2, line 12-15) and Results (page 9, line 20 - page10, line 8).

5. *The cut-off point of 2 vFDG-positive regions in Figure 2B and 2C seems arbitrary when looking at Figure 2A. Because distinguishing patients with CK > 2000 from those with less or between patients with histological score > 3 from those with less does not contribute to the accurate diagnosis of PM/DM, these comparisons can be excluded from the paper. Instead, Authors might be interested in finding a feasible vFDG-based method to distinguish between patients with PM/DM and those with ALS.*
   We did not found correlation between numbers of vFDG-positive regions and CK in simple regression analysis. However, there was a mild correlation between these parameters. We think Figure 2B is necessary in this study.
   We replaced figure 2C with a figure showing correlation between the number of vFDG-positive regions and the total histological scores.
   To test whether vFDG-based methods distinguish PM/DM and ALS strictly, with a blinded manner, is quite difficult. Because ALS patients have muscle atrophy often begins asymmetrically, which can be easily identified by FDG PET with CT. We calculated SUV in PM/DM and ALS and found FDG uptake was significantly higher in PM/DM than that in ALS. We also confirmed a strong correlation between SUV and vFDG. Thus it is feasible that vFDG can distinguish PM/DM from ALS. A prospective study in a future may is necessary.

6. *Because the total vFDG score in each muscle contains the information on severity of inflammation, not only the number of vFDG-positive regions, but the total sum of the total; vFDG scores in all muscles in individual patients can be worth analyzing for its usefulness.*
   According to the reviewer's comment, we also analyzed using total vFDG score using both mediatinum blood vessels and liver as positivity criteria. We added sentences in Discussion (page 13 , line 8 - 13).

7. *Please be more specific in describing the methods of determining the vFDG score. Do blood vessels in the mediastinum only mean the thoracic aorta? Or other vessels such as vena cava as well? Which dud Authors compare the maximum or average FDG uptake? or just subject impression?*

vFDG is a method which depends on visual impression. Mediastinum blood vessels (pool) is widely used as a reference and it includes aorta as well as vena cava (J Clin Oncl 2007;25:571).

8. *The proportion of patients who underwent FDG PET without CT scan needs to be provided. Without CT images, FDG uptakes in individual muscle regions must be challenging. This point needs to be discussed as another limitation of this study.*
   We described the number of the patients underwent FDG PET without CT scan (page 9, line 2). FDG uptake in the muscle regions which we selected for the study can be evaluated by axial slices of FDG PET without CT scan. We described this point in the limitation of the study (page 3, line 9).

Minor concerns:

1. *As Authors separately scored the inflammation and the fiber necrosis/regeneration of muscle biopsy specimens, it might be interesting to see the correlation between each score and the FDG uptake in the muscle where muscle biopsy was performed. Although this analysis is not directly related to the purpose of this study, it may give novel insight into the interpretation of FDG accumulation in myositis.*
   This is a very important point. We analyzed correlation between each histological score and the FDG uptake in the muscle of biopsy site. We found correlation between FDG uptake and scores of necrosis/regeneration and inflammation. The score of necrosis/regenelation correlates more strongly with FDG uptake. We added sentences in Methods (page 8, line 8, 11) and Results (page 10 , line 6 - 8, 15 - 23).

2. *The possibility that the small proportion of patients who underwent FDG PET after receiving treatment affected the analyses and made the results less impressive can be discussed.*
   We added description about this point in Discussion (page 15, line 8 - 9).

3. *Figure 4 is not clearly reproduced in the PDF file.*
   We improved clarity of Figure 4.

Reviewer: 2

Major points:

1. *In this article, FDG uptake was visually evaluated (visually identified FDG uptake, vFDG). This method might be practically reasonable for the screening of yet-unidentified cancer in the whole body; however, one of the major objectives in this paper is the detection of inflammation in the skeletal muscle. The percentage of vFDG-positive PM/DM patients, 63.6%, is not so high of authors are handling the patients before treatment. The rationale to place SUV = 2 as a positivity criterion for vFDG should be clarified in detail.*
   One of the disadvantages of visual assessment in FDG PET is the limitation on available positive criterion. Mediastinum blood vessels (SUV is around 2) are generally used for positive criterion. Because FDG uptake in PM/DM muscle is relatively moderate (mean SUVmax was apporoximately 1.4), low sensitivity is inevitable when mediastinum blood vessels is used as the positive criterion. An investigation for more suitable positive criterion or grading scales is necessary in a future study.

2. *Authors found that mean SUV max was significantly higher in patients with PM/DM patients than*

in those with ALS. This results is reasonable and potentially interest, but, if FDG-PET meaningful in the clinical practice, the comparison of these values should be done in patients with non-inflammatory myopathies including muscular dystrophies and metabolic myopathies and authors should try to persuade the readers that FDG-PET is useful in excluding non-inflammatory myopathies. This is the most important, because authors should answer the following question: why the physicians add such expensive study like PET in the initial evaluation of PM/DM?
We agree with the reviewer that a comparison between PM/DM and non-inflammatory myopathies is necessary. However, this is a retrospective study and there were only few patients with muscular dystrophy who underwent FDG PET for detection of occult cancers. This is a limitation of our study (page 3, line 10-11, page 15, line 10 - 11). We added sentences in Discussion (page15, line20 - 21).

3. *Figure 3 is beautiful and impressive carrying the impression of the differences between MRI-detected lesions and PET-positive lesions. However, I need to present two arguments: firstly, vFDG positive lesion is actually artificial and the FDG+ lesion can be enlarged or shrinked according to the window setting. For example, in this figure, Hamstrings do not show increased FDG uptake but how about the real value of SUV max? Is this not evaluated as compared with normal muscles? Secondary, the authors stated that affected muscles showed the "diffuse" high signal in MRI. This is definitely not true. Furthermore, "localized" vFDG lesion is actually dependent upon the window setting. These sentences should be rewritten.*
We confirmed the FDG-positive lesion not only in FDG PET/CT image but also in FDG PET image (direct image of FDG uptake). Figure 3 demonstrates difference in images of inflammatory muscle lesions represented by different modalities. The most noteworthy point is that FDG uptake is more obvious within the muscle than the surface. The finding is quite different from that of MRI image. We made changes in Abstract (page 2, line 17-19), Results (page 11, line 12 - 13), Discussion (page 13 , line 14 - 22), and the legend of the Figure 3.

4. *The background of increased SUV in the skeletal muscle should be clarified in detail. Authors quoted refs. 8 and 9 and stated that accumulation of FDG is caused by active cells such as macrophages, young granulation tissues, and fibroblasts; however, the pathologic picture of acute phase of PM/DM is different from those of RA and authors should describe what kind of cells/structures contribute to the accumulation of FDG in PM/DM. How is the contribution of degeneration /regeneration muscle cells?*
This is an important point, however, to the best of our knowledge, there are no reports about background of FDG uptake in myositis. As same as other cell types, muscle cells in regeneration stage consume glucose and may increase FDG uptake. In fact, we found a strong correlation between FDG uptake (both mean SUVmax and vFDG score) and the score of necrosis/regeneration. We made changes in Discussion (page 12 , line 10 - 17).

5. *It is an overstatement that "FDG uptake was confined to metabolically active cells".*
We made changes in page 13, line 21 - 22 .

Minor points

1. *Make sure that what criteria are used to diagnose PM/DM based on muscle pathology or simply the presence of cutaneous findings.*
We added sentences in page 6, line 21 – page 7, line 2.

2. *One of the drawbacks of this study is the original design of this study: cancer screening of PM/DM patients. Authors proposed that scans should include the entire length of thigh for the future screening of PM/DM. How about the lower legs?*
   It is the best to scan all of the body including lower legs. However, the scan time become considerably prolonged. Thus we propose that entire length of thigh is required at least for PM/DM screening in the clinical practice.

3. *Quadriceps muscle is composed of four muscles. These four are of course indistinguishable in the clinical practice, but this study deals with radiological /pathological discussions. Please specify.*
   We did not discriminate the individual muscles of quadriceps in this study protocol. It is beyond the scope of our current study.

4. *Some English usages are awkward and difficult to interpret. Authors are recommended to be checked by somebody whose tongue is English and who can understand the content of this article thoroughly.*
   English of our manuscript was checked by a native speaker well versed in science.

Reviewer: 3

1. *No definitions have been given for PM or DM. Furthermore, there is subdivision of DM into Definite and Probable without definition.*
   We used criteria proposed by Bohan et al (ref. 15). In addition, we classified the patients using recent clinico-pathological criteria proposed in 119th European Neuromuscular Centre (ENMC) international workshop (ref. 18). We added sentences in page 5, line 23 – page 6, line 2 and page 9, line 7-8.

2. *The last sentence of p4 makes no sense – Muscle biopsies were conducted in all patients in whom inflammatory infiltrates and muscle fiber necrosis and /or expression of HLA class 1 on muscle fibers were observed.*
   We made corrections in this sentence (page 5, line 23 – page 6, line 2).

3. *The next sentence implies that not all biopsies showed inflammation – A few patients showed either inflammatory infiltrates or muscle fiber necrosis. How does this fit in with their definition of cases?*
   These patients classified as probable PM/DM by Bohan's criteria. Today, we can examine muscle pathology more precisely with immunostaining. If an overexpression of HLA class 1 molecules on the muscle fibers were found, we included the patients in the study. We also carefully excluded other diseases. We made corrections in this part (page 5, line 23 - page 6, line 2).

4. *vFDG was positive in 21/33 patients, and MRI in 20/25 implying MRI is more sensitive. Did any MRI negative patients show FDG positivity.*
   We found there were no patient with MRI negative and FDG positive (Table 1). MRI may be more sensitive. However, MRI was not evaluated in a blinded manner in this study. Comparison was done only in MRI examined regions. Sensitivity between MRI and FDG PET should be compared

in prospective study in a future. We made changes in Results (page 11, line 5 - 6, page 13, line 24 - page 14, line 5).

5. *They imply that FDG may be more specific than MRI but don't' have the evidence to support that. Certainly MRI appearances maybe similar in myositis and other myopathies, but they have no data about the specificity of FDG.*
   We would like to mention different backgrounds of MRI and FDG PET. We do not think FDG PET is more specific.

6. *Relating to that, ALS is a poor control group, being a neurogenic rather than myopathic disorder. They should also look at those with conditions that cause diagnostic difficulties, such as limb-girdle dystrophies.*
   The study was retrospective and we could not find adequate number of patients with muscular dystrophy underwent FDG PET. We wrote this point in the limitation of the study (page 3, line 10 - 11).