AJNR

Comparison of [18F] FDG-PET/MRI and Clinical Findings for Assessment of Suspected Lumbar Facet Joint Pain: A Prospective Study to Characterize Candidate Nonanatomic Imaging Biomarkers and Potential Impact on Management

V.T. Lehman, F.E. Diehn, S.M. Broski, M.A. Nathan, B.J. Kemp, N.B. Larson, R.A. Shelerud, J.S. Brault, M.P. Halasy and T.P. Maus

Comparison of [¹⁸F] FDG-PET/MRI and Clinical Findings for Assessment of Suspected Lumbar Facet Joint Pain: A Prospective Study to Characterize Candidate Nonanatomic Imaging Biomarkers and Potential Impact on Management

V.T. Lehman, F.E. Diehn, S.M. Broski, M.A. Nathan, B.J. Kemp, N.B. Larson, R.A. Shelerud, J.S. Brault, M.P. Halasy, and T.P. Maus

ABSTRACT

BACKGROUND AND PURPOSE: Prior retrospective studies have suggested that both T2 hyperintensity and gadolinium enhancement on fat-suppressed MR imaging are associated with lumbar facet joint pain, but prospective evaluation of FDG-PET/MR imaging with a standardized protocol and correlation to clinical findings are lacking. The primary aim was to prospectively assess a standardized FDG-PET/MRI protocol in patients with suspected facetogenic low back pain, with determination of the concordance of imaging and clinical findings.

MATERIALS AND METHODS: Ten patients with clinically suspected facetogenic low back pain were prospectively recruited with a designation of specific facet joints implicated clinically. Subsequently, patients underwent an FDG-PET/MR imaging examination with gadolinium. Each facet joint was graded for perifacet signal change on MR imaging and FDG activity. The frequency and correlation of MR imaging, FDG-PET, and clinical findings were determined.

RESULTS: FDG activity showed high concordance with high overall MR imaging scores (concordance correlation coefficient = 0.79). There was concordance of the clinical side of pain with the side of high overall MR imaging scores and increased FDG activity on 12/20 (60%) sides. Both a high overall MR imaging score (concordance correlation coefficient = 0.12) and FDG-PET findings positive for increased activity (concordance correlation coefficient = 0.10) had low concordance with the specific clinically implicated facet joints. Increased FDG activity or high MR imaging scores or both were present in only 10/29 (34%) facet joints that had been clinically selected for percutaneous intervention. Eleven (11%) facet joints that had not been selected for treatment demonstrated these imaging findings.

CONCLUSIONS: There was low concordance of perifacet signal change and FDG activity with clinically implicated facet joints. This could indicate either the potential to change patient management or a lack of biomarker accuracy. Therefore, additional larger randomized studies with the use of comparative medial branch blocks would be useful to further investigate the clinical utility of these findings.

ABBREVIATIONS: MBB = medial branch block; ρ_{CCC} = concordance correlation coefficient; SUVmax = standard uptake value maximum

Facet joints have been implicated in 15%–45% of cases of low back pain, but diagnosis remains challenging.[1] Anatomic imaging findings of facet joint arthropathy or clinical findings are not considered reliable indicators of individual lumbar facet joints to target for treatment.[1,2] Current standards for identification of a painful facet joint require sequential blinded comparative medial branch blocks (MBBs) with local anesthetic and >80% pain relief.[3] While percutaneous treatments exist, more effective and more durable options are needed for many patients. Investigation of new treatment agents and modalities is already underway.[4,5] Thus, identification of nonanatomic facet joint imaging biomarkers to facilitate diagnosis, direct individualized target-specific treatment, and help assess response to existing and experimental treatments is highly desirable.

Recently, combined PET/MR imaging units have become clinically available, enabling simultaneous technique assessment that minimizes differences due to temporal fluctuation of findings or anatomic misregistration. Several retrospective

studies have suggested that facet/perifacet findings on fat-suppressed MR imaging are more prevalent in patients with low back pain, but prospective clinical investigations with standardized correlation of imaging findings to specifically implicated facet joints or MBBs have not been performed to evaluate and fully characterize these findings.[6-10] [18F] FDG is the most widely available PET radiotracer, has been used to assess inflammation in a variety of conditions, allows quantitative uptake analysis, and has the potential to reflect inflammatory facet joint arthropathy.[11] Such facet joint FDG activity has been described in a prevalence study, a study of the cervical spine, and a case report in the lumbar spine,[12] but there is little prospective information addressing lumbar facet joints.[13,14]

The primary aim of this pilot study was to prospectively assess a standardized FDG-PET/MR imaging protocol with gadolinium in patients with suspected facetogenic axial low back pain. This includes assessment of the concordance of nonanatomic MR imaging FDG-PET and clinical findings on a patient and facet joint level. A secondary aim was to obtain concordance, prevalence, and descriptive information of these potential biomarkers for planning of large future studies using comparative MBBs.

MATERIALS AND METHODS
Patient Selection and Initial Clinical Evaluation
Institutional review board approval was obtained for this Health Insurance Portability and Accountability Act–compliant study. This study was registered under clinicaltrials.gov (NCT02921490). Subjects were identified primarily or referred to a spine or pain clinic and evaluated by a pain or spine specialist. Patients with unilateral or bilateral axial low back pain were assessed for inclusion criteria: 1) minimum 60% likelihood of facet joint–related axial low back pain based on clinical history, pain location and character, tenderness to palpation, and positive facet joint loading maneuvers; 2) no substantial pain or tenderness below the iliac crest/sacroiliac joint region or predominant pain radiating to the lower extremities; and 3) clinically a 60%–80% or 81%–100% confidence that pain was facetogenic for each side of the back with pain. These percentages were determined by the overall clinical impression using all available history and physical examination information based on the clinician's cumulative clinical experience compared with all other patients he or she had ever evaluated. Also documented were the response to facet joint loading maneuvers and tenderness to palpation on each side; the duration and severity of pain, with the latter being rated on the numeric rating scale; and the presence of known inflammatory arthropathy.

Exclusion criteria included younger than 50 years of age, a procedure within the preceding 2 months (lumbar spine injection, lumbar spine surgery, or facet joint intervention [MBB, radiofrequency ablation, or steroid injection]), history of major trauma to the lumbar spine, metastatic malignancy, conditions with increased radiosensitivity, pregnancy, compression fracture, and contraindication to MR imaging.

Before FDG-PET/MR imaging, the specialist clinician indicated which specific facet joints she or he was planning to target with percutaneous treatment (either steroid injection or MBB and radiofrequency ablation) on the basis of clinical evaluation alone as a representation of implicated facet joints. Treatment based on PET/MR imaging results was not mandated, and the clinical response to any injection was not recorded. Because the PET/MR imaging examination was performed before any treatment, clinicians were ultimately free to recommend interventions based on the initial plan and/or PET/MR imaging results using their best clinical judgment.

Imaging Protocol
Images were acquired on a fully integrated simultaneous 3T PET/MR imaging scanner (Signa PET/MR imaging; GE Healthcare, Milwaukee, Wisconsin) with TOF capability. Weight, height, and blood glucose levels were recorded for all patients. All patients fasted for >4 hours, had a blood glucose level of <200 mg/dL, and were injected with 10 mCi of [18F] FDG (mean, 10.4 mCi). PET acquisition was limited to a single bed position centered over the lumbar spine. The PET scan was acquired for 15 minutes after an uptake period of 60 minutes (mean, 62 minutes). PET images were reconstructed with 3D ordered subset expectation maximization with TOF, 2 iterations, 28 subsets, and a 5-mm Gaussian postfilter into a 192×192 matrix covering a 50-cm FOV. PET attenuation correction was performed using a 3D dual-echo radiofrequency spoiled gradient recalled-echo sequence; scatter, randoms, deadtime, and decay corrections were also applied. MR imaging sequences obtained of the lumbar spine and acquired simultaneously with the PET scan included sagittal T1, sagittal T2, sagittal T2 with chemical fat saturation, axial T1, and axial T2 with chemical fat saturation without gadolinium. After gadolinium administration, axial and sagittal postgadolinium T1-weighted images were obtained with chemical fat saturation. These axial images covered the L1–L2 through L5–S1 facet joints with a straight axial pack. All sequences were performed as 2D fast spin-echo. Specific scan parameters are provided in On-line Table 1.

MR Imaging Analysis
Both radiologists interpreting MR imaging (T.P.M., F.E.D.) have extensive experience in spine imaging interpretation and pain management intervention, with Certificates of Added Qualification in neuroradiology. MR imaging readers were blinded to FDG-PET data and all clinical data. Images were viewed anonymously on an Advantage Workstation (GE Healthcare). MR images of each facet joint were scored twice, first using all available noncontrast MR images (including T2-weighted fat-suppressed images) and then a second time with the addition of gadolinium contrast-enhanced sequences.

Each lumbar facet joint L1–L2 through L5–S1 was individually graded for the following features: osseous signal change/enhancement and soft-tissue perifacet edema/enhancement. The grading scale represented a more granular modification of that used by Czervionke and Fenton,[6] which combined perifacet and bone features into a single grading scale.

Osseous signal change was graded on a I–III scale for T2 hyperintensity or enhancement: grade 0 (normal), none; grade I, present in 1 articular process; grade II, present in both articular

processes; and grade III, extending into the pedicle, transverse process, or lamina.

Soft-tissue perifacet signal change was graded on a I–IV scale for edema and enhancement: grade 0 was normal; grade I (minimal), thin and curvilinear confined to the posterior facet joint capsule without further extension; grade II (mild), in soft tissue extending beyond the facet joint capsule encompassing <50% of the facet joint perimeter; grade III (moderate to high), in soft tissue beyond the facet joint capsule encompassing 51%–100% of the facet joint perimeter; and grade IV (high), extending into the neural foramen, ligamentum flavum, and/or radial extension equal or >1 cm from the facet joint margin in any direction.

Each instance of a discrepant score between the 2 MR imaging readers was resolved with a tie-breaking grading by a third reader with a Certificate of Added Qualification in neuroradiology (V.T.L.), who was also blinded to all clinical and FDG-PET information at the time of interpretation. The tie-breaking radiologist was instructed to choose 1 of the 2 original scores.

FDG-PET Imaging Analysis

FDG-PET/MR images were rated independently by 2 radiologists, one with a Certificate of Added Qualification in nuclear medicine (S.M.B.) and one with double board certification by the American Board of Radiology and American Board of Nuclear Medicine (M.A.N.). MIM software was used for image review (MIM Software, Cleveland, Ohio). Both unfused PET and fused PET/MR images were available for review, but only the noncontrast sagittal and axial T1-weighted images without fat saturation were available for image coregistration. The readers remained blinded to all clinical information.

The readers graded the FDG activity subjectively on a 0–III scale: 0 = normal, I = mildly increased, II = moderately increased, or III = markedly increased activity. Consensus grades for locations with an initial discrepancy were determined during a second review by both nuclear medicine radiologists. Additionally, the standard uptake value maximum (SUVmax) was determined at each facet joint by drawing a VOI to encompass the entire osseous facet joint. The SUVmax of the bone marrow within the L3 vertebral body and the blood pool activity within the abdominal aorta at the L3 level were both determined by placement of a 1-cm-diameter ROI on axial images. In addition to the facet joint SUVmax, ratios normalized by values from the L3 vertebral body and aorta were also considered for analysis.

Data Analysis

MR Imaging and FDG-PET Data.
To facilitate comparison with FDG and clinical findings, we combined the perifacet and osseous grades into an overall MR imaging score: normal (bone or perifacet grade of 0); low MR imaging score (bone grade of I–II or perifacet grade I–II); or high MR imaging score (bone grade III or soft-tissue perifacet grade of III–IV). The highest of the consensus grades (osseous T2 signal, soft-tissue perifacet T2 signal, osseous enhancement, or soft-tissue perifacet enhancement) was used for this overall MR imaging score. To facilitate comparison with MR imaging and clinical findings, we assigned each joint an overall FDG-PET score of either normal or increased (grades I–III) FDG activity.

The effect of gadolinium on grade and score assignments was determined by the rate at which these differed from the addition of gadolinium-enhanced images. The rates of high overall score designation on the basis of perifacet grade versus osseous grade were also determined.

Comparison of the Clinical Side of Pain and Implicated Facet Joints with Imaging Results.
Comparison of clinical findings with imaging results had 2 main components: First, the imaging findings were evaluated for concordance with clinical findings on a patient side (right or left). For this evaluation, any side with pain that also had at least 1 facet joint with high-grade MR imaging findings and/or increased FDG activity was considered concordant, regardless of specific facet joint levels. Similarly, the absence of both pain and these imaging findings was considered concordant. Sides with a high overall MR imaging score and/or increased FDG activity without pain or pain without such imaging findings were considered discordant. The rationale was that determination of the precise level of facet joint pain clinically is thought to be difficult, but pain generators typically produce ipsilateral rather than contralateral pain.[2]

Statistical Analysis.
Binary and categoric variables (eg, PET grade) were summarized as counts and percentages, while continuous measures were summarized by means and SDs or medians and interquartile ranges. Distributional assumptions for continuous-valued traits were assessed, and appropriate transformations were considered, as necessary. All analyses were conducted using the R statistical and computing software (http://www.r-project.org/), and statistical significance was declared at an α level of .05.

Concordance measures were used to quantify agreement among overall high MR imaging scores, increased FDG activity, and facet joints that had been selected for treatment on the basis of clinical evaluation. To account for the multiple lumbar levels per given patient, we applied the repeated-measures concordance correlation coefficient (ρ_{CCC}) using the *cccrm* R package[15] (https://www.rdocumentation.org/packages/cccrm/versions/1.2.1) to generate point estimates and 95% confidence intervals.

We tested whether the SUV measurements significantly differed by positive PET grade level using likelihood ratio tests based on a random-intercept linear mixed-model for each of the 3 SUV measurements. The reduced model used the dichotomized PET grade coding (0 versus 1–3), whereas the full model considered each positive grade as a separate factor level. For each model, SUV measurements were log-transformed to satisfy assumptions of normality.

We quantified discrimination of dichotomous PET grade coding by each SUV measurement using area under the receiver operating characteristic curve statistics. We applied the pooled repeated-measures approach[16] in the cvAUC R package (https://cran.r-project.org/web/packages/cvAUC/index.html) to account for within-patient correlation structure, providing point estimates and corresponding 95% CIs.

Given the ordinal nature of the grades, the ordinal Krippendorf α was used to determine the interrater reliability. The 95% confidence interval was estimated using the grouped

bootstrap to account for within-patient clustering, and the intervals were constructed using the percentile approach.

RESULTS

Participant Demographics, Lumbar Enumeration, and Clinical Characteristics

The study cohort consisted of 10 subjects including 7 (70%) women, with a mean age of 63 years (range, 50–79 years). This allowed evaluation of 100 facet joints on 20 sides (left or right) of the lumbar spine. Pain duration was >12 months (9/10 subjects, 90%) and 6–12 months (1/10, 10%). One (10%) subject had been diagnosed with underlying undifferentiated inflammatory arthropathy (patient 1).

The mean severity of pain on the numeric rating scale was 5 (range, 3–8) at the time of clinical evaluation. Four (40%) patients had unilateral pain, whereas 6 (60%) had bilateral pain. Of the 16 sides with pain, the clinical confidence that the pain was due to facet joint origin before imaging was in the 60%–80% range for 8 (50%) sides and 81%–100% range for 8 (50%) sides. Clinical features are detailed in On-line Table 2.

MR Imaging Scores

On MR imaging, 21 (21%) of 100 facet joints demonstrated a high overall MR imaging score, more frequently due to soft-tissue perifacet findings rather than osseous findings, specifically, 2/21 (10%) on the basis of osseous findings (grade III/III), 12/21 (57%) on the basis of soft-tissue perifacet findings (grades III–IV/IV), and 7/21 (33%) due to both perifacet and osseous findings. Of the 19 facet joints with high overall scores and grade III–IV soft tissue perifacet signal, 13 (68%) had a grade of III, and 6 (32%) had a grade of IV.

The use of gadolinium increased the number of facet joints with high overall MR imaging scores, predominantly due to an increased grade of soft-tissue perifacet signal change. Specifically, 10/21 (48%) joints were upgraded to a high MR imaging score with the addition of gadolinium-enhanced images compared with T2 fat-saturated images alone, whereas 11/21 (52%) were assigned a high overall MR imaging score with both T2-weighted and gadolinium-enhanced images. One of the 10 (10%) upgraded facet joints was on the basis of osseous enhancement, whereas 9/10 (90%) were upgraded on the basis of soft-tissue perifacet enhancement.

In the subset of 19 facet joints with high overall MR imaging scores due to grade III–IV soft-tissue perifacet signal change with gadolinium, the perifacet signal grades on fat-suppressed T2-weighted images were variable but frequently low-grade with the following frequencies: grade 0 ($n = 1$); grade I ($n = 5$); grade II ($n = 7$); grade III ($n = 5$); and grade IV ($n = 1$). Eight of 9 (89%) joints with high-grade bone signal (grade III) had the finding on both T2-weighted and gadolinium-enhanced images, whereas 1/9 (11%) had high-grade change on gadolinium-enhanced images alone.

Fifty-three of 100 (53%) facet joints were designated normal on T2-weighted images, but only 10 (10%) were scored as normal on both T2-weighted fat-suppressed and gadolinium-enhanced images (On-line Fig 1). Sixty-nine (69%) facet joints had a low overall MR imaging score. This includes 43/69 (62%) joints that were normal on T2 fat-saturated images alone but demonstrated low-grade findings on gadolinium-enhanced images due to the presence of low-grade capsular enhancement with or without mild perifacet extension. The other 26/69 (38%) had low-grade findings on both T2 fat-suppressed images and gadolinium-enhanced images.

FDG-PET Imaging Scores

On PET, 17/100 (17%) facet joints demonstrated increased FDG activity overall scores, 10 (10%) low-grade (grade I) and 7 (7%) moderate-to-high grade (grade II–III). All 17 (100%) FDG-positive facet joints demonstrated a high MR imaging overall score. Two patients (20%) had no (0%) facet joints with increased FDG activity; these were also the only 2 patients without high-grade MR imaging change. Increased FDG activity was highly correlated with the presence of a high overall MR imaging score within a specific facet joint (ρ_{CCC} = 0.78; 95% CI, 0.68–0.85). Furthermore, 12/12 (100%) sides with facet joints with high overall MR imaging scores also had increased FDG activity. Comparisons of the clinical features and major imaging findings are presented in On-line Tables 2 and 3.

The median (interquartile range) SUVmax for the 83 facet joints visually graded as 0 was 1.50 (0.55), compared with 1.85 (1.50) for the 17 facet joints with increased FDG activity (grades I–III). Comparison of SUVmax, SUVmax/aorta ratio, and SUVmax/L3 vertebral body ratios also demonstrated a trend of increasing SUVmax values with visually assigned grades ($P < .001$ for all analyses). There were also instances of overlapping values, in particular between grade 0 and grade I. The SUVmax values considered here include only a VOI of the osseous facet joint itself for consistency and reproducibility, whereas the visually assigned scores also considered perifacet soft-tissue and other osseous activity if present.

The area under the curve values for dichotomous discrimination between normal and abnormal (any grade) facet joints using the SUVmax, SUVmax/aorta ratio, or SUVmax/L3 vertebral body ratio were relatively comparable given overlapping CIs, with the SUVmax/L3 vertebral body ratio corresponding to the highest point estimate. Specifically, the area under the curve values were 0.72 (95% CI, 0.57–0.88) for SUVmax, 0.84 (95% CI, 0.73–0.96) for SUVmax/aorta, and 0.87 (95% CI, 0.74–0.99) for the SUVmax/L3 vertebral body.

Interrater Reliability

The interrater reliability of overall MR imaging scores was moderate, with an α estimate of 0.77 (95% CI, 0.52–0.92). The interrater reliability of the PET scores was high, with an α estimate of 0.85 (95% CI, 0.65–0.98).

Concordance of Facet Joint Imaging Findings to Side of Pain and Specific Facet Joints Implicated Clinically

There was concordance of the clinical side of pain, the side of the high overall MR imaging score, and the side of increased FDG activity on 12/20 (60%) sides (On-line Table 3). This includes 10 sides (50%) with imaging findings positive for concordance and 2 sides (10%) with imaging findings negative for concordance and the absence of suspected facet joint pain.
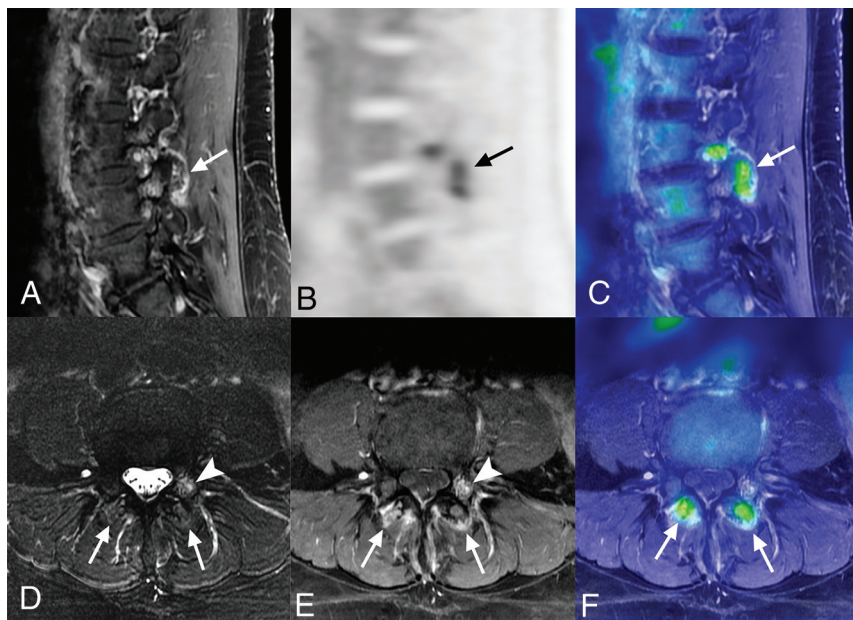
FIGURE. Clinical concordance of the sides of pain and imaging findings, but discordance of specific implicated facet joints. Clinically, this patient had bilateral low back pain and had been prescribed bilateral L4–L5 and L5–S1 facet joint injections. There were high-grade MR imaging scores and increased FDG activity of the bilateral L3–L4 facet joints, but not of the bilateral L4–L5 or L5–S1 facet joints. Sagittal fat-suppressed T1-weighted image with gadolinium demonstrates high-grade perifacet enhancement of the left L3–L4 facet joint (*arrow*), but not of the L4–L5 or L5–S1 facet joints (*A*). Sagittal FDG-PET and fused PET/MR images (*B* and *C*) also demonstrate increased L3–L4 perifacet FDG activity (*arrows*). Note that the FDG activity is distinguishable from areas of vascular enhancement but is not visually increased in the area of L3 pedicle enhancement (*arrowhead in D and E*). The perifacet signal change of the bilateral L3–L4 facet joints on axial fat-suppressed T2-weighted images is identified (*arrows, D*) but is more apparent on axial fat-suppressed T1-weighted images with gadolinium (*arrows, E*). An axial fused PET/MR image also demonstrates increased perifacet FDG activity of the bilateral L3–L4 facet joints (*arrows, F*).

Eight (40%) sides were discordant, including 5/20 (25%) sides with pain without a high overall MR imaging score or increased FDG activity and 3/20 (15%) sides with such imaging findings and no clinically suspected facet joint pain.

Tenderness to palpation was reported on 14/20 (70%) sides. Nine of 14 (64%) sides with tenderness had ipsilateral high overall MR imaging scores or increased FDG activity, while 5/14 (36%) sides with tenderness lacked these imaging findings. Conversely, a high overall MR imaging score and FDG activity were identified in 3/6 (50%) sides without tenderness to palpation.

Facet joints that had been selected for treatment before imaging demonstrated low correlation with those with either a high overall MR imaging score (ρ_{CCC} = 0.12; 95% CI, 0.002–0.23) or increased FDG activity (ρ_{CCC} = 0.10; 95% CI, −0.04–0.24). Treatment directed toward 29/100 (29%) facet joints was indicated on initial clinical evaluation before the PET/MR imaging. Increased FDG activity or high overall MR imaging scores or both were present in 10 of these 29 (34%) facet joints. Additionally, increased FDG activity or a high MR imaging score or both were present in 11 facet joints that were not initially prescribed treatment. Therefore, 10/21 (48%) facet joints with high overall MR imaging scores and/or FDG activity corresponded to facet joints specifically planned for targeted percutaneous intervention.

Moreover, specific facet joints with high-grade overall MR imaging scores and/or increased FDG activity completely differed from those facet joints originally prescribed treatment in 4/10 (40%) subjects, partially overlapped in 5/10 (50%) subjects, and completely corresponded in only 1 (10%) subject. A comparison of the facet joints originally selected for treatment with those with increased FDG activity or high-grade MR imaging scores is presented in On-line Table 3. Examples comparing the sides and specific facet joints implicated clinically with those demonstrating positive imaging findings are provided in the Figure and On-line Fig 2.

The most marked consensus MR imaging and FDG findings (grades II–III FDG findings, grade III osseous MR imaging signal, grade IV soft-tissue perifacet signal) were less frequent but were found on sides with concordant pain. Of the subgroup of 7 facet joints that had moderate ($n = 4$) or high-grade ($n = 3$) FDG activity, all 7 (100%) were on sides ($n = 7$ sides) with clinically concordant pain. Of the subset of 9 facet joints that had high-grade (grade III) osseous MR imaging signal, all 9 (100%) were on a side ($n = 5$ sides) of clinically concordant pain. Of the subset of 6 facet joints that had high-grade (grade IV) perifacet signal, all 6 (100%) were on a side ($n = 6$ sides) of clinical pain.

DISCUSSION

The primary result of this study was a low concordance between perifacet signal change or FDG activity with clinically implicated facet joints. This finding could indicate the potential of imaging findings to substantially change management if these prove to be predictive of facet joint pain. Alternatively, this low concordance could indicate a lack of sensitivity or specificity of imaging findings. In either case, further investigation seems warranted to facilitate appropriate interpretation of imaging findings and assess the utility (or lack thereof) for selection of joints for treatment. Most important, future study would benefit from the use of dual or triple comparative MBBs (multiple injections) to support/refute facet joint pain because a single intervention such as a steroid injection has an unacceptably high placebo rate and is not considered the standard of diagnosis in the pain medicine community.[3] As with this study, prior lumbar spine facet joint studies with MR imaging or FDG-PET have not incorporated such comparative MBBs.[6,9,10,13] Unlike the current study, these prior studies have not incorporated a standardized imaging protocol with

prospective clinical evaluation. We believe it will be useful if future investigations with multiple interventions are predicated on pilot data with standardized imaging protocols and image scoring such as in this study. Indeed, our results help provide justification for and facilitate planning of future investigations.

Specifically, the high correlation of FDG activity with MR imaging signal indicates that these findings are likely close surrogates for initial assessment, and the use of a single technique (MR imaging or FDG-PET) may be sufficient. However, comparison of MR imaging and FDG findings as possible biomarkers of treatment response would require additional study with FDG-PET/MR imaging at multiple time points. While not well-studied in the facet joints, there is evidence that FDG activity can indicate an early response to systemic treatment in other inflammatory conditions, including inflammatory arthritis.[17,18] We found that high-grade perifacet findings were more common with gadolinium, supporting the assertion that further investigation with MR imaging protocols including gadolinium, rather than fat-suppressed T2 images alone, would be reasonable. The results also provide initial assessment of the approximate frequency and concordance of potential imaging biomarkers on multiplane fat-suppressed T2, fat-suppressed T1 gadolinium-enhanced, and FDG-PET images using a standardized protocol.

In many instances, imaging findings were concordant with the side of pain but indicate partially or completely different facet joints for targeted treatment in nearly every patient and in more than half of all implicated facet joints. Complete concordance of imaging and clinical findings and clinical confidence would indicate that use of imaging biomarkers is unlikely to change the targets for treatment. Conversely, complete discordance of the imaging and clinical findings would raise doubt that the candidate biomarkers actually indicate facet joint pain at all. If either of these scenarios were observed, the utility of future investigation of these imaging findings could be questioned.

The instances of positive imaging findings in the absence of ipsilateral pain and vice versa raise the possibility that these candidate biomarkers may have limited specificity and sensitivity. The lack of imaging findings on the side of clinically suspected facet joint pain in some instances could either represent a limitation of imaging biomarker sensitivity or could potentially indicate non-facet joint origin of pain.

Because the facet joints with the most marked MR imaging signal and FDG activity were always correlated with the side of pain, future studies will ideally be large enough to determine the significance of this subset. Low-grade perifacet enhancement was nearly ubiquitous, however, raising caution of overinterpretation of this finding in clinical practice or research studies.

Unlike in the current study, Czervionke and Fenton[6] retrospectively reported that back and/or leg pain was always present on the side of unilateral high-grade facet joint signal change, perhaps due to methodologic differences of study design, clinical assessment, and standardization of imaging protocol. The correlation of back pain and imaging findings was not otherwise reported, and clinically implicated facet joints were not determined.[6] Sawicki et al[14] reported that FDG-PET/MR imaging–directed cervical facet joint steroid

injections were associated with improved clinical pain scores relative to clinically selected facet joints in patients without imaging findings positive for increased activity, though MBBs were not used and gadolinium was not administered. They also found that MR imaging T2 hyperintensity and FDG activity may be surrogates.[14]

One prior study suggested that FDG-PET/MR imaging findings of the lumbar spine may help identify various causes of sciatica, including inflamed facet joints.[19] Limited evidence indicates that muscular or nerve FDG activity may be present in a painful lower extremity.[19,20] Future studies can build on these results and those of the current study to investigate the utility of FDG-PET/MR imaging in assessing other causes of axial and radicular lumbar pain.

This study has some limitations. The number of patients was small, and the study was performed at a single institution. However, the degree of concordance of MR imaging signal change and FDG was high enough and the lack of concordance between facet joints originally implicated with the imaging findings was marked enough to be statistically salient. Recruitment of patients specifically being evaluated by back pain specialists could introduce selection bias and could limit generalization of the results to other patient populations. Furthermore, each patient was evaluated by only a single clinician and subjective assignment of the likelihood of facet joint origin of pain, even with standardized items on the clinical examination, could potentially differ among clinicians. Finally, invasive tests such as MBBs to provide more definitive evidence to support or refute facet joint origin of pain were not used in this pilot study.

CONCLUSIONS

Although the most marked imaging findings were always found on a side with pain, there was low concordance of perifacet signal change and FDG activity with clinically implicated facet joints overall. This could indicate either a potential to change patient management or a lack of biomarker accuracy. Therefore, additional larger randomized studies with the use of comparative MBBs are needed to clarify the clinical utility of these findings. There was a high concordance of MR imaging and FDG-PET findings, suggesting that a single technique (MR imaging or FDG-PET) design may be reasonable for future studies.

ACKNOWLEDGMENTS

REFERENCES

1. Schwarzer AC, Aprill CN, Derby R. **Clinical features of patients with pain stemming from the lumbar zygapophysial joints: is the lumbar facet syndrome a clinical entity?** *Spine (Phila Pa 1976)* 1994;19:1132–37 Medline

2. McCall IW, Park WM, O'Brien JP. **Induced pain referral from posterior lumbar elements in normal subjects.** *Spine (Phila Pa 1976)* 1979;4:441–46 Medline

3. Bogduk N. *ISIS Practice Guidelines for Spinal Diagnostic and Treatment Procedures.* 2nd ed. International Spine Intervention Society, San Francisco; 2014

4. Weeks EM, Platt MW, Gedroyc W. **MRI-guided focused ultrasound (MRgFUS) to treat facet joint osteoarthritis low back pain–case series of an innovative new technique.** *Eur Radiol* 2012;22:2822–35 CrossRef Medline

5. Wu J, Zhou J, Liu C, et al. **A prospective study comparing platelet-rich plasma and local anesthetic (LA)/corticosteroid in intra-articular injection for the treatment of lumbar facet joint syndrome.** *Pain Pract* 2017;17:914–24 CrossRef Medline

6. Czervionke LF, Fenton DS. **Fat-saturated MR imaging in the detection of inflammatory facet arthropathy (facet synovitis) in the lumbar spine.** *Pain Med* 2008;9:400–06 CrossRef Medline

7. D'Aprile P, Tarantino A, Lorusso V, et al. **Fat saturation technique and gadolinium in MRI of lumbar spinal degenerative disease.** *Neuroradiol J* 2006;19:654–71 CrossRef Medline

8. Friedrich KM, Nemec S, Peloschek P, et al. **The prevalence of lumbar facet joint edema in patients with low back pain.** *Skeletal Radiol* 2007;36:755–60 CrossRef Medline

9. Lakadamyali H, Tarhan NC, Ergun T, et al. **STIR sequence for depiction of degenerative changes in posterior stabilizing elements in patients with lower back pain.** *AJR Am J Roentgenol* 2008;191:973–79 CrossRef Medline

10. Suri P, Dharamsi AS, Gaviola G, et al. **Are facet joint bone marrow lesions and other facet joint features associated with low back pain? A pilot study.** *PM R* 2013;5:194–200 CrossRef Medline

11. Wu C, Li F, Niu G, et al. **PET imaging of inflammation biomarkers.** *Theranostics* 2013;3:448–66 CrossRef Medline

12. Houseni M, Chamroonrat W, Zhuang H, et al. **Facet joint arthropathy demonstrated on FDG PET.** *Clin Nucl Med* 2006;31:418–19 Medline

13. Rosen RS, Fayad L, Wahl RL. **Increased 18F-FDG uptake in degenerative disease of the spine: characterization with 18F-FDG PET/CT.** *J Nucl Med* 2006;47:1274–80 Medline

14. Sawicki LM, Schaarschmidt BM, Heusch P, et al. **Value of (18) F-FDG PET/MRI for the outcome of CT-guided facet block therapy in cervical facet syndrome: initial results.** *J Med Imaging Radiat Oncol* 2017;61:327–33 CrossRef Medline

15. Carrasco JL, Phillips BR, Puig-Martinez J, et al. **Estimation of the concordance correlation coefficient for repeated measures using SAS and R.** *Comput Methods Programs Biomed* 2013;109:293–304 CrossRef Medline

16. LeDell E, Petersen ML, van der Laan MJ. **Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates.** *Electron J Stat* 2015;9:1583–1607 CrossRef Medline

17. Fosse P, Kaiser MJ, Namur G, et al. **(18)F- FDG PET/CT joint assessment of early therapeutic response in rheumatoid arthritis patients treated with rituximab.** *Eur J Hybrid Imaging* 2018;2:6 CrossRef Medline

18. Nguyen BJ, Burt A, Baldassare RL, et al. **The prognostic and diagnostic value of 18F-FDG PET/CT for assessment of symptomatic osteoarthritis.** *Nucl Med Commun* 2018;39:699–706 CrossRef Medline

19. Cipriano PW, Yoon D, Gandhi H, et al. **(18)F-FDG PET/MRI in chronic sciatica: early results revealing spinal and nonspinal abnormalities.** *J Nucl Med* 2018;59:967–72 CrossRef Medline

20. Biswal S, Behera D, Yoon DH, et al. **[18F]FDG PET/MRI of patients with chronic pain alters management: early experience.** *EJNMMI Phys* 2015;2:(Suppl 1):A84 CrossRef Medline