Reliability of change in lumbar MRI findings

Poster No.: C-0539  
Congress: ECR 2012  
Type: Scientific Exhibit  
Authors: L. Berg¹, O. Gjertsen², C. Hellum², G. F. Neckelmann¹, L. G. Johnsen³, G. E. Eide¹, A. Espeland¹; ¹Bergen/NO, ²Oslo/NO, ³Trondheim/NO  
Keywords: Arthritides, Observer performance, MR, Spine, Neuroradiology spine, Musculoskeletal spine  
DOI: 10.1594/ecr2012/C-0539

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method is strictly prohibited.

You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.

www.myESR.org
Purpose

- Little reliability data exist regarding change in magnetic resonance imaging (MRI) findings over time.
- In daily clinical practice direct comparison of new and old images is usually the preferred method to rate change in MRI findings.
- To our knowledge, only one study (of Modic changes) has examined the reliability of change in lumbar spine MRI findings over time [1]. Based on their finding of moderate reliability by direct comparison of images, the authors of that study recommended non-comparison when studying the course of Modic changes, i.e. independent evaluation of the two sets of images.
- Our aim was to evaluate the reliability of change in lumbar MRI findings both by non-comparison and by direct comparison of images.

Methods and Materials

Patients and images

- We included 126 (of 173) chronic low back pain patients (mean age 41.6 years; 61 men, 65 women) who were randomized to disc prosthesis surgery or non-surgical treatment [2] and had both pre-treatment and a 2-year follow-up lumbar spine MRI available.
- MRI was obtained at different trial sites using different protocols and typically included 1.5 T sagittal T1- and T2-weighted images and (mostly T2-weighted) axial images of L3/L4, L4/L5 and L5/S1.

Image evaluation

- At L3/L4, L4/L5 and L5/S1, two experienced radiologists independently evaluated progress (included new findings) and regress of MRI findings from pre-treatment to the 2-year follow-up, both by non-comparison and by direct comparison of images.
- Rating of Modic changes [1,3-5], posterior high intensity zone (HIZ) in the disc [6,7], nucleus pulposus signal [8], disc height reduction evaluated subjectively [9-11], measured disc height decrease [12], disc contour [5] and facet arthropathy (FA) [13] was based on existing rating criteria and published illustrations from the Spine Pain Outcomes Research Trial [7].

Statistical analysis
• Prevalence- and bias-adjusted kappa (PABAK) for interobserver agreement was calculated for progress and regress of each MRI finding at each disc level.
• The impact of image evaluation method on PABAK was assessed using generalized estimating equations.

Results

• Both observers evaluated FA at all levels, and other findings at 125-126 L3/L4 levels, 87-88 of 89 non-operated L4/L5 levels and 72-73 of 74 non-operated L5/S1 levels, depending on the finding. Missing data in one to two cases were due to artifacts from an adjacent disc prosthesis or disagreement on level due to transitional vertebrae.
• By comparison of images interobserver agreement was good on progress and regress (PABAK 0.63-1.00) for Modic changes, posterior HIZ, disc height, and disc contour at L3-S1 and for nucleus pulposus signal and FA at L3/L4; and moderate (PABAK 0.46-0.59) on decreasing nucleus signal and increasing FA at L4-S1.
• PABAK values were significantly higher by direct comparison of new and old images than by non-comparison (p<0.001).
• Fig. 1 on page 3, Fig. 2 on page 4 and Fig. 3 on page 5 illustrate agreements and disagreements.

Images for this section:
Fig. 1: Pre-treatment (a#d) and 2-year follow-up (e#h) MRI of one patient. T2-weighted images (b#d, f#h) are shown in the order of patient's left to right, T1-weighted images in a and e corresponding to T2-weighted images in b and f, respectively. At L4/L5, two observers A and B agreed on conversion of Modic changes superior to disc from primary type I (a#d; arrowhead) to primary type II (e#h) and on unchanged grey nucleus pulposus signal on T2-weighted images (d, h). A and B agreed on an unchanged rating by comparison of images for posterior HIZ at L4/L5 (unchanged no HIZ and unchanged HIZ), but disagreed by non-comparison (unchanged no HIZ, progress from no HIZ to HIZ) (b#d, f#h). At L5/S1, A and B agreed on progress from no HIZ (c, d) to HIZ (g, h)
**Fig. 2:** Pre-treatment (a#d) and 2-year follow-up (e#h) MRI of one patient treated with disc prosthesis in L5/S1; sagittal T2-weighted images (a, e), axial proton density-weighted images at L4/L5 (b, c) and L5/S1 (d), and corresponding axial T2-weighted follow-up images at L4/L5 (f, g) and L5/S1 (h). Image planes shown in b#d and f#h are marked on a and e (broken lines). At L4/L5, two observers A and B agreed on unchanged normal disc contour (b#c, f#g) and stable disc height (unchanged slight reduction and unchanged no reduction) (a, e). At L4/L5, A and B agreed on unchanged nucleus pulposus signal (a, e) by comparison of images (unchanged grey, unchanged bright), but disagreed by non-comparison (progress from bright to grey, regress from grey to bright). By comparison, A and B agreed on FA at L4/L5 (unchanged slight FA, unchanged no FA) (b#c, f#g), and on unchanged no FA at L5/S1 (d, h). By non-comparison, A and B disagreed on FA both at L4/L5 (regress from moderate to slight FA, unchanged no FA) and at L5/S1 (regress from slight to no FA, unchanged no FA)
Fig. 3: Fig 3 T1- (a, c) and T2-weighted (b,d) MRI pre-treatment (a,b) and at 2-year follow-up (c, d) in one patient treated with disc prosthesis in L5/S1. Adjacent level degeneration is an important issue in patients treated with disc prosthesis or fusion surgery. In this patient readers A and B agreed on no adjacent level degeneration at 2-year follow-up (unchanged no Modic or disc findings at L4/L5).
Conclusion

- For changes in lumbar MRI findings over time, direct comparison of new and old images can provide moderate or good interobserver agreement, and better agreement than non-comparison.

- The present results support the clinical practice of comparing images to assess change.

- Further studies are needed to confirm our findings and to assess the clinical relevance of change in lumbar MRI findings over time.

References


Personal Information

Linda Berg MD.

Department of Radiology, Haukeland University Hospital and
Section for Radiology, Department of Surgical Sciences, University of Bergen, Bergen, Norway.

E-mail: Linda.Berg@kir.uib.no