

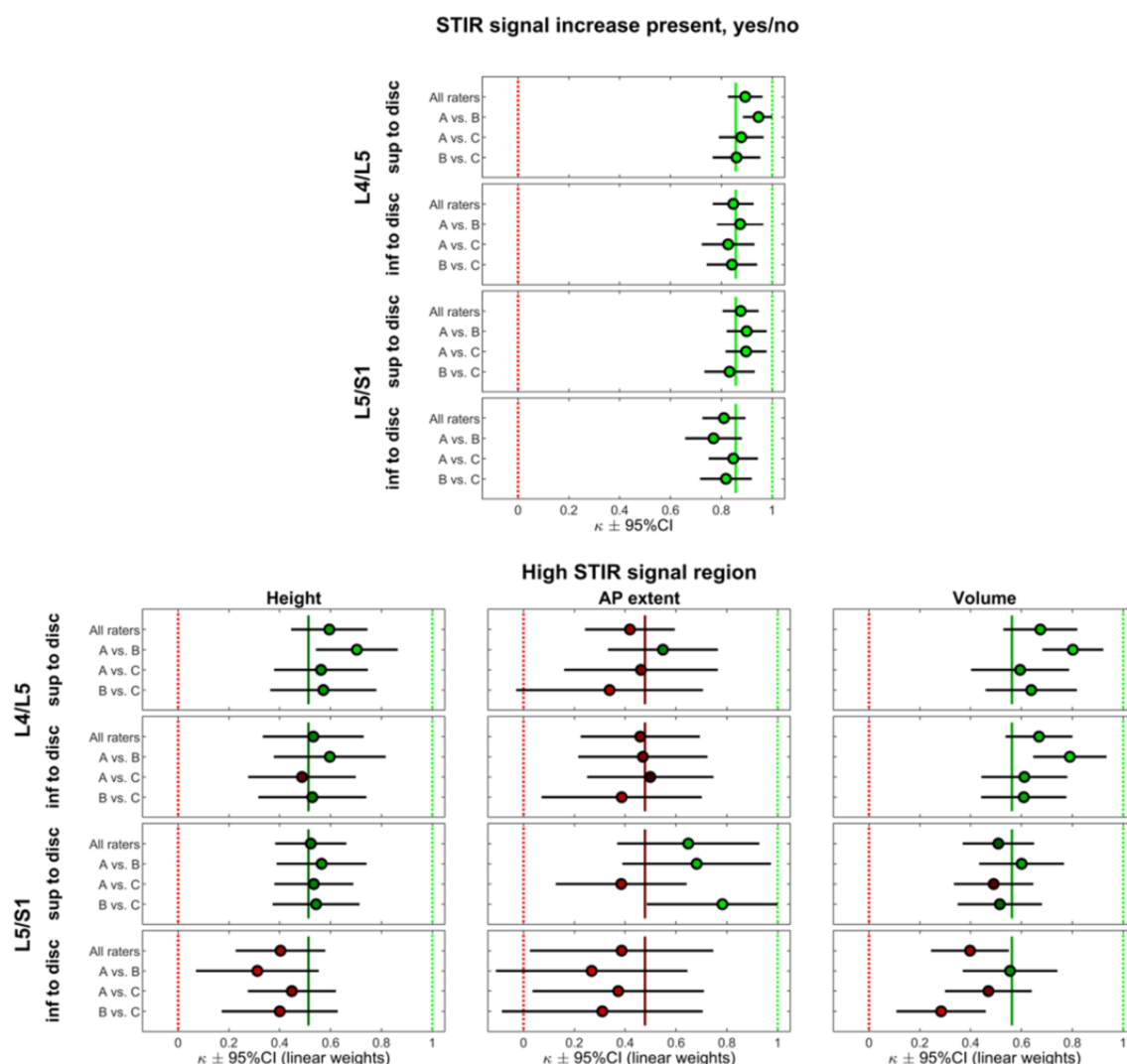
Appendix:

Lumbar spine evaluations of Modic changes with STIR – are they reliable?

Table A1 Inclusion and exclusion criteria for the [REDACTED] -trial

Inclusion criteria	Exclusion criteria
<ul style="list-style-type: none">• Age between 18 and 65 years• LBP of more than 6 months duration in the area below the 12th rib and above the gluteal folds with a Numerical Rating Scale (NRS) pain intensity score of ≥ 5 (mean of three 0–10 NRSs: current LBP, the worst LBP within the last 2 weeks, and the usual/ mean LBP within the last 2 weeks)• MRI-confirmed lumbar disc herniation within the preceding 2 years• Type I and/or type II MC in the vertebral body marrow at the same level as the previously herniated disc. For patients with previous surgery for disc herniation, the MC has to be located at an operated level• Written informed consent	<ul style="list-style-type: none">• Allergy to [REDACTED]• Allergy/hypersensitivity to any of the excipients of the study drug• Current pregnancy or lactation• Kidney (creatinine) or hepatic (ALAT/ASAT) laboratory values above the normal range• Phenylketonuria (Følling's disease)• Mononucleosis or leukaemia• Any specific diagnosis that may explain the patient's low back symptoms (e.g., tumour, fracture, spondyloarthritis, infection, spinal stenosis)• Previous low back surgery (L1–S1) for reasons other than disc herniation (e.g., fusion, decompression, disc prosthesis)• Surgery for disc herniation within the last 12 months• Previous surgery for disc herniation, but MC located at level(s) that has/have not been operated on only• Reservation about the intake of gelatine (the capsules used to encapsulate the study medicine contains gelatine, which, among other things, is produced using ingredients derived from pigs)• Regular use of glucocorticoids• Regular use of opioids with the exception of codeine and tramadol• Not understanding [REDACTED] language• Unlikely to adhere to treatment and/or complete follow-up (e.g., serious ongoing psychiatric disease, drug abuse, plans to move)• Antibiotic treatment within the preceding one month before treatment start• Contraindications to MRI (e.g., cardiac pacemaker electrodes, metal implant in the eye or brain, claustrophobia)• Unwilling to participate

Fig. A1 Categorical STIR variables: forest plot for kappa values with 95% confidence intervals, including pairwise results



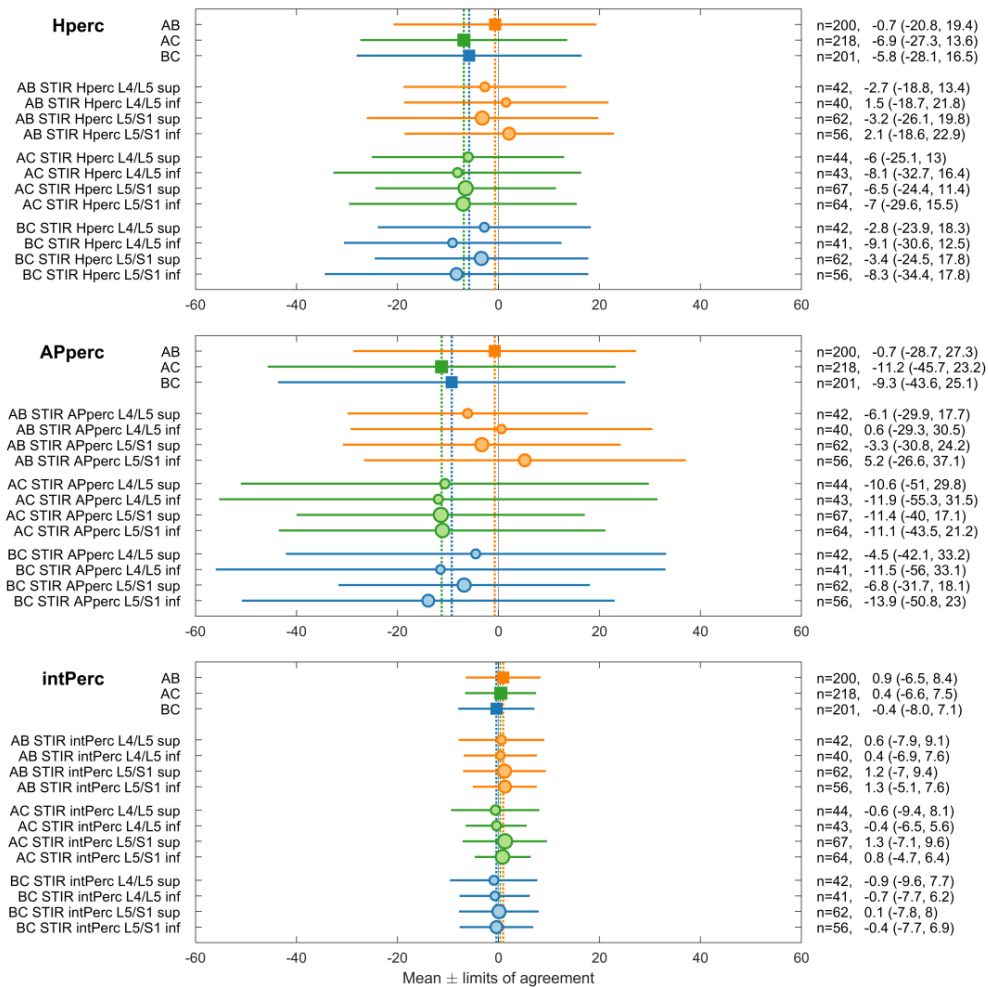
The figure shows kappa values with 95% confidence intervals (CIs) for all observers and for observer pairs AB, AC and BC for variables describing Modic change (MC) related STIR signal increases superior (sup) and inferior (inf) to the L4/L5 and L5/S1 discs. These variables were presence (yes/no), height (four categories), anteroposterior (AP) extent (three categories) and volume (four categories) of region with high STIR signal. Mean Fleiss' kappa value for agreement between all raters across all four levels L4-S1 is marked with a bold vertical line. This line and circles representing kappa values are green for kappa values higher than 0.50 (the midpoint of the moderate agreement category) and otherwise red.

Table A2 Categorical STIR variables: kappa values with 95% confidence intervals

	Pairwise (Cohen's kappa)						All Raters (Fleiss' kappa)		
	A vs. B		A vs. C		B vs. C		N	Kappa (95%CI)	Mean
	N	Kappa (95%CI)	N	Kappa (95%CI)	N	Kappa (95%CI)			
STIR signal increase (MC related) present, yes/no									0.86
L4/L5 sup to disc	120	0.95 (0.89,1.00)	120	0.88 (0.79,0.97)	120	0.86 (0.77,0.95)	120	0.89 (0.83,0.96)	
L4/L5 inf to disc	120	0.87 (0.78,0.97)	120	0.83 (0.72,0.93)	120	0.84 (0.74,0.94)	120	0.85 (0.77,0.93)	
L5/S1 sup to disc	120	0.90 (0.82,0.98)	120	0.90 (0.82,0.98)	120	0.83 (0.73,0.93)	120	0.88 (0.81,0.95)	
L5/S1 inf to disc	120	0.77 (0.66,0.88)	120	0.85 (0.75,0.94)	120	0.82 (0.72,0.92)	120	0.81 (0.73,0.89)	
STIR signal increase only inside MC, yes/no									0.30
L4/L5 sup to disc	42	0.44 (0.15,0.74)	44	0.29 (0.06,0.52)	42	0.28 (0.06,0.49)	42	0.28 (0.06,0.50)	
L4/L5 inf to disc	40	0.40 (0.10,0.71)	43	0.22 (-0.03,0.47)	41	0.17 (-0.03,0.36)	40	0.19 (-0.03,0.42)	
L5/S1 sup to disc	62	0.64 (0.43,0.86)	67	0.34 (0.16,0.51)	62	0.36 (0.17,0.56)	62	0.41 (0.23,0.59)	
L5/S1 inf to disc	56	0.48 (0.24,0.71)	64	0.23 (0.03,0.44)	56	0.33 (0.16,0.50)	56	0.31 (0.12,0.50)	
Height of the region with high STIR signal, four categories									0.51
L4/L5 sup to disc	42	0.70 (0.54,0.86)	44	0.56 (0.38,0.75)	42	0.57 (0.36,0.78)	42	0.59 (0.44,0.74)	
L4/L5 inf to disc	40	0.60 (0.38,0.82)	43	0.49 (0.28,0.70)	41	0.53 (0.32,0.74)	40	0.53 (0.33,0.73)	
L5/S1 sup to disc	62	0.56 (0.39,0.74)	67	0.53 (0.38,0.69)	62	0.54 (0.37,0.71)	62	0.52 (0.38,0.66)	
L5/S1 inf to disc	56	0.31 (0.07,0.55)	64	0.45 (0.27,0.62)	56	0.40 (0.17,0.63)	56	0.40 (0.23,0.58)	
AP extent of the high STIR signal, three categories									0.48
L4/L5 sup to disc	42	0.55 (0.33,0.76)	44	0.46 (0.16,0.76)	42	0.34 (-0.03,0.71)	42	0.42 (0.24,0.60)	
L4/L5 inf to disc	40	0.47 (0.22,0.72)	43	0.50 (0.25,0.75)	41	0.39 (0.07,0.70)	40	0.46 (0.22,0.69)	
L5/S1 sup to disc	62	0.68 (0.39,0.97)	67	0.38 (0.13,0.64)	62	0.78 (0.49,1.00)	62	0.65 (0.37,0.93)	
L5/S1 inf to disc	56	0.27 (-0.11,0.64)	64	0.37 (0.04,0.71)	56	0.31 (-0.09,0.71)	56	0.39 (0.03,0.75)	
Volume of the high STIR signal, four categories									0.56
L4/L5 sup to disc	42	0.80 (0.68,0.92)	44	0.59 (0.40,0.79)	42	0.64 (0.46,0.82)	42	0.67 (0.53,0.82)	
L4/L5 inf to disc	40	0.79 (0.65,0.93)	43	0.61 (0.44,0.78)	41	0.61 (0.44,0.78)	40	0.67 (0.54,0.80)	
L5/S1 sup to disc	62	0.60 (0.43,0.77)	67	0.49 (0.33,0.65)	62	0.51 (0.35,0.68)	62	0.51 (0.37,0.65)	
L5/S1 inf to disc	56	0.56 (0.37,0.74)	64	0.47 (0.30,0.64)	56	0.28 (0.11,0.46)	56	0.40 (0.24,0.55)	

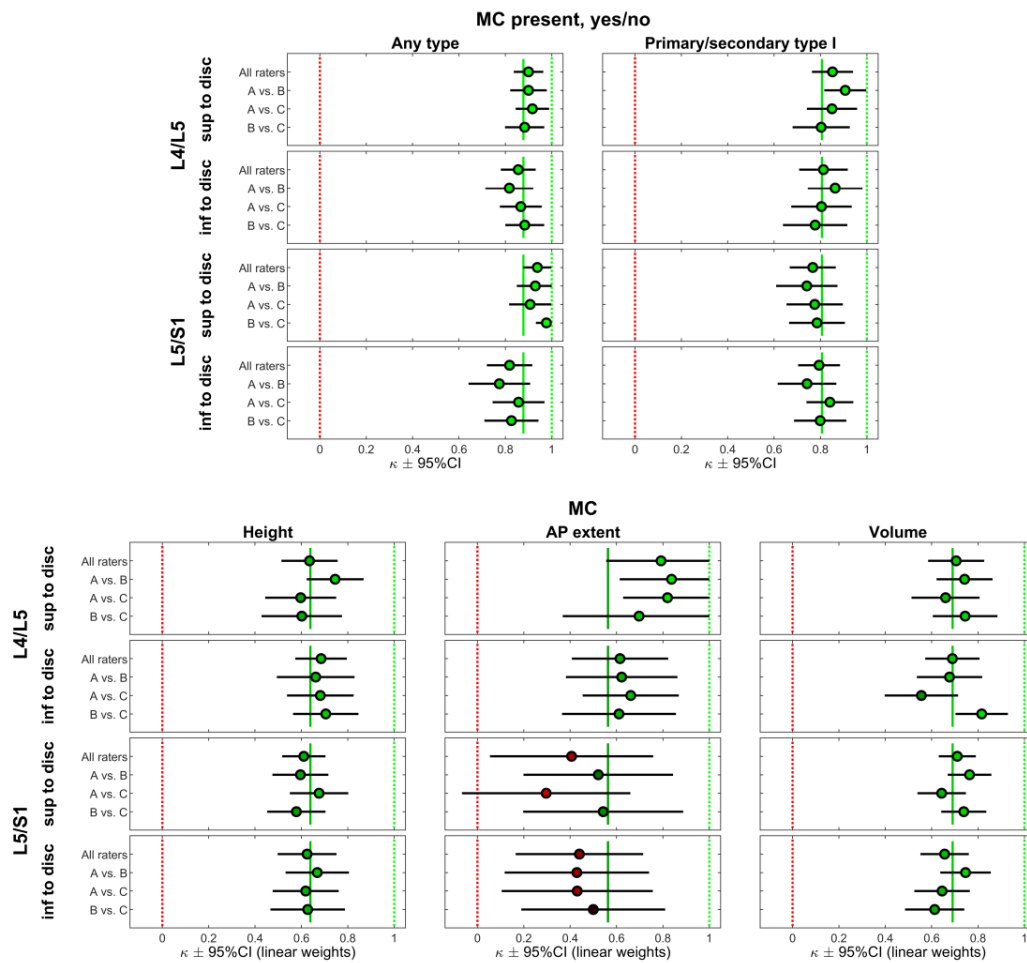
STIR= short tau inversion recovery, N=number of patients with data from both (or all) observers, CI=confidence interval, MC=Modic change, sup=superior, inf=inferior
 Tabled are kappa values with 95% CIs for observer pairs AB, AC and BC and for all observers for variables describing MC related STIR signal increases superior (sup) and inferior (inf) to the L4/L5 and L5/S1 discs. For variables with more than two categories, linearly weighted kappa values are presented

Fig. A2 Numerical STIR variables: forest plot for means of differences and limits of agreement, including results for individual endplates



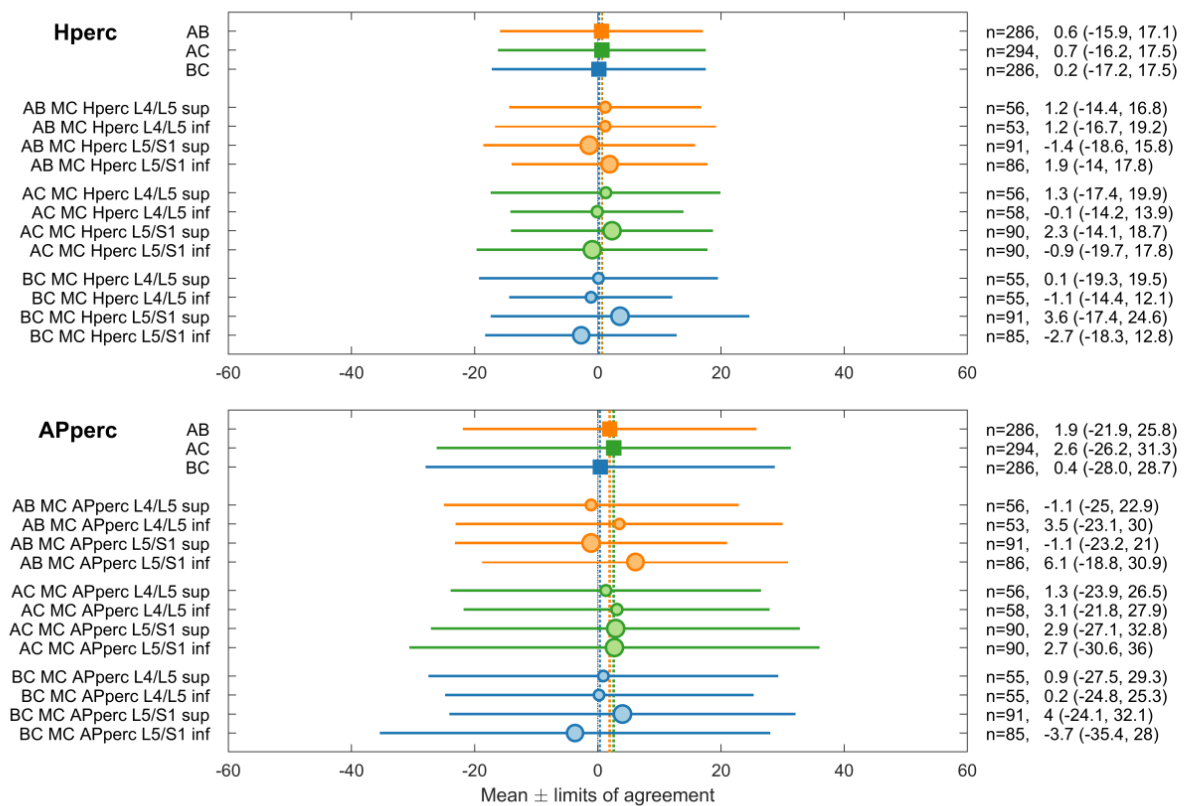
The figure shows mean of differences with 95 % limits of agreement in observer pairs AB, AC and BC for three numerical variables describing STIR signal increases related to Modic changes. Each variable was evaluated at four endplates (superior and inferior to the L4/L5 and L5/S1 discs). For each variable, means for all endplates are first displayed followed by means for individual endplates. Values are % points. Hperc denotes height of region with high STIR signal in % of the height of the vertebral body marrow; APperc denotes anteroposterior (AP) extent of the high STIR signal in % of the midsagittal AP diameter of the endplate; intPerc denotes maximum intensity of the STIR signal in % on a scale from normal vertebral body marrow intensity (0 %) to CSF intensity (100 %).

Fig. A3 Categorical MC variables on T1/T2: forest plot for kappa values with 95% confidence intervals, including pairwise results



The figure shows kappa values with 95% confidence intervals (CIs) for all observers and for observer pairs AB, AC and BC for variables describing Modic changes (MCs) superior (sup) and inferior (inf) to the L4/L5 and L5/S1 discs on T1/T2 weighted fast spin echo images. These variables were presence of any type of MCs (yes/no), presence of primary or secondary type I MCs (yes/no), and height (four categories), anteroposterior (AP) extent (three categories) and volume (four categories) of the MCs. Mean Fleiss' kappa value for agreement between all raters across all four levels L4-S1 is marked with a bold vertical line. This line and circles representing kappa values are green for kappa values higher than 0.50 (the midpoint of the moderate agreement category) and otherwise red.

Fig. A4 Numerical MC variables on T1/T2: forest plot for means of differences and limits of agreement, including results for individual endplates



The figure shows mean of differences with 95 % limits of agreement in observer pairs AB, AC and BC for two numerical variables describing Modic changes (MCs) on T1/T2 weighted fast spin echo images. Each variable was evaluated at four endplates (superior and inferior to the L4/L5 and L5/S1 discs). For each variable, means for all endplates are first displayed followed by means for individual endplates. Values are % points. Hperc denotes height of the MC in % of the height of the vertebral body marrow; APperc means anteroposterior (AP) extent of the MC in % of the midsagittal AP diameter of the endplate.

Table A3 Categorical MC variables: kappa values with 95% confidence intervals

	Pairwise (Cohen's kappa)						All Raters (Fleiss' kappa)		
	A vs. B		A vs. C		B vs. C		N	Kappa (95%CI)	Mean
	N	Kappa (95%CI)	N	Kappa (95%CI)	N	Kappa (95%CI)			
MC of any type and size present, yes/no									0.88
L4/L5 sup to disc	120	0.90 (0.82,0.98)	120	0.92 (0.84,0.99)	120	0.88 (0.80,0.97)	120	0.90 (0.84,0.96)	
L4/L5 inf to disc	120	0.82 (0.71,0.92)	120	0.87 (0.78,0.96)	120	0.88 (0.80,0.97)	120	0.86 (0.78,0.93)	
L5/S1 sup to disc	120	0.93 (0.85,1.00)	120	0.91 (0.82,1.00)	120	0.98 (0.93,1.00)	120	0.94 (0.88,1.00)	
L5/S1 inf to disc	120	0.77 (0.64,0.91)	120	0.86 (0.74,0.97)	120	0.83 (0.71,0.94)	120	0.82 (0.72,0.92)	
Primary or secondary type 1 MC present, yes/no									0.81
L4/L5 sup to disc	120	0.91 (0.82,1.00)	120	0.85 (0.74,0.96)	120	0.80 (0.68,0.93)	120	0.85 (0.76,0.94)	
L4/L5 inf to disc	120	0.86 (0.75,0.98)	120	0.80 (0.67,0.93)	120	0.78 (0.64,0.92)	120	0.81 (0.71,0.92)	
L5/S1 sup to disc	120	0.74 (0.61,0.87)	120	0.78 (0.65,0.90)	120	0.78 (0.66,0.91)	120	0.77 (0.67,0.87)	
L5/S1 inf to disc	120	0.74 (0.62,0.87)	120	0.84 (0.74,0.94)	120	0.80 (0.69,0.91)	120	0.79 (0.70,0.88)	
Height of MC, four categories									0.64
L4/L5 sup to disc	56	0.75 (0.62,0.87)	56	0.60 (0.44,0.75)	55	0.60 (0.43,0.78)	54	0.63 (0.51,0.76)	
L4/L5 inf to disc	53	0.66 (0.49,0.83)	58	0.68 (0.54,0.83)	55	0.70 (0.56,0.85)	53	0.68 (0.57,0.80)	
L5/S1 sup to disc	91	0.60 (0.48,0.72)	90	0.68 (0.55,0.80)	91	0.58 (0.45,0.70)	90	0.61 (0.52,0.70)	
L5/S1 inf to disc	86	0.67 (0.53,0.80)	90	0.62 (0.48,0.76)	85	0.63 (0.47,0.79)	85	0.62 (0.50,0.75)	
AP extent of MC, three categories									0.56
L4/L5 sup to disc	56	0.84 (0.61,1.00)	56	0.82 (0.63,1.00)	55	0.70 (0.37,1.00)	54	0.79 (0.56,1.00)	
L4/L5 inf to disc	53	0.62 (0.38,0.86)	58	0.66 (0.45,0.87)	55	0.61 (0.36,0.86)	53	0.61 (0.41,0.82)	
L5/S1 sup to disc	90	0.52 (0.20,0.84)	89	0.30 (-0.07,0.66)	90	0.54 (0.20,0.89)	89	0.41 (0.05,0.76)	
L5/S1 inf to disc	86	0.43 (0.12,0.74)	90	0.43 (0.10,0.76)	85	0.50 (0.19,0.81)	85	0.44 (0.16,0.71)	
Volume of MC, four categories									0.69
L4/L5 sup to disc	56	0.74 (0.62,0.86)	56	0.66 (0.51,0.81)	55	0.74 (0.60,0.88)	54	0.71 (0.58,0.83)	
L4/L5 inf to disc	53	0.68 (0.54,0.82)	58	0.56 (0.40,0.71)	55	0.82 (0.70,0.93)	53	0.69 (0.57,0.81)	
L5/S1 sup to disc	91	0.76 (0.67,0.86)	90	0.64 (0.54,0.75)	91	0.74 (0.64,0.84)	90	0.71 (0.63,0.79)	
L5/S1 inf to disc	86	0.75 (0.64,0.86)	90	0.64 (0.53,0.76)	85	0.61 (0.49,0.74)	85	0.66 (0.55,0.76)	

MC=Modic change, N=number of patients with data from both (or all) observers, CI=confidence interval, sup=superior, inf=inferior

Tabled are kappa values with 95% CIs for observer pairs AB, AC and BC and for all observers for variables describing MCs superior (sup) and inferior (inf) to the L4/L5 and L5/S1 discs. For variables with more than two categories, linearly weighted kappa values are presented