

Supplementary Table 1 Analysis of intra-observer reliability

	Evaluation 1 ¹	Evaluation 2 ¹	Difference ¹	p-value ²	Correlation coefficient (p-value) ³
DD score	2.7 (\pm 1.0)	2.8 (\pm 1.0)	-0.12 (\pm 0.33)	0.083	0.886 (<0.001)
EP score	1.8 (\pm 1.7)	1.9 (\pm 1.7)	-0.10 (\pm 0.30)	0.025	0.890 (<0.001)
FD score	1.6 (\pm 0.9)	1.6 (\pm 0.9)	-0.04 (\pm 0.28)	0.317	0.965 (<0.001)
FS score	0.4 (\pm 0.7)	0.5 (\pm 0.8)	-0.10 (\pm 0.42)	0.096	0.790 (<0.001)

¹Mean (\pm SD)

²Statistical significance of difference between evaluation 1 and evaluation 2, Wilcoxon signed-rank test

³Spearman correlation coefficient

DD score – Disc degeneration score, EP score – Endplate score, FD score – Facet degeneration score, FS score – Foraminal stenosis score

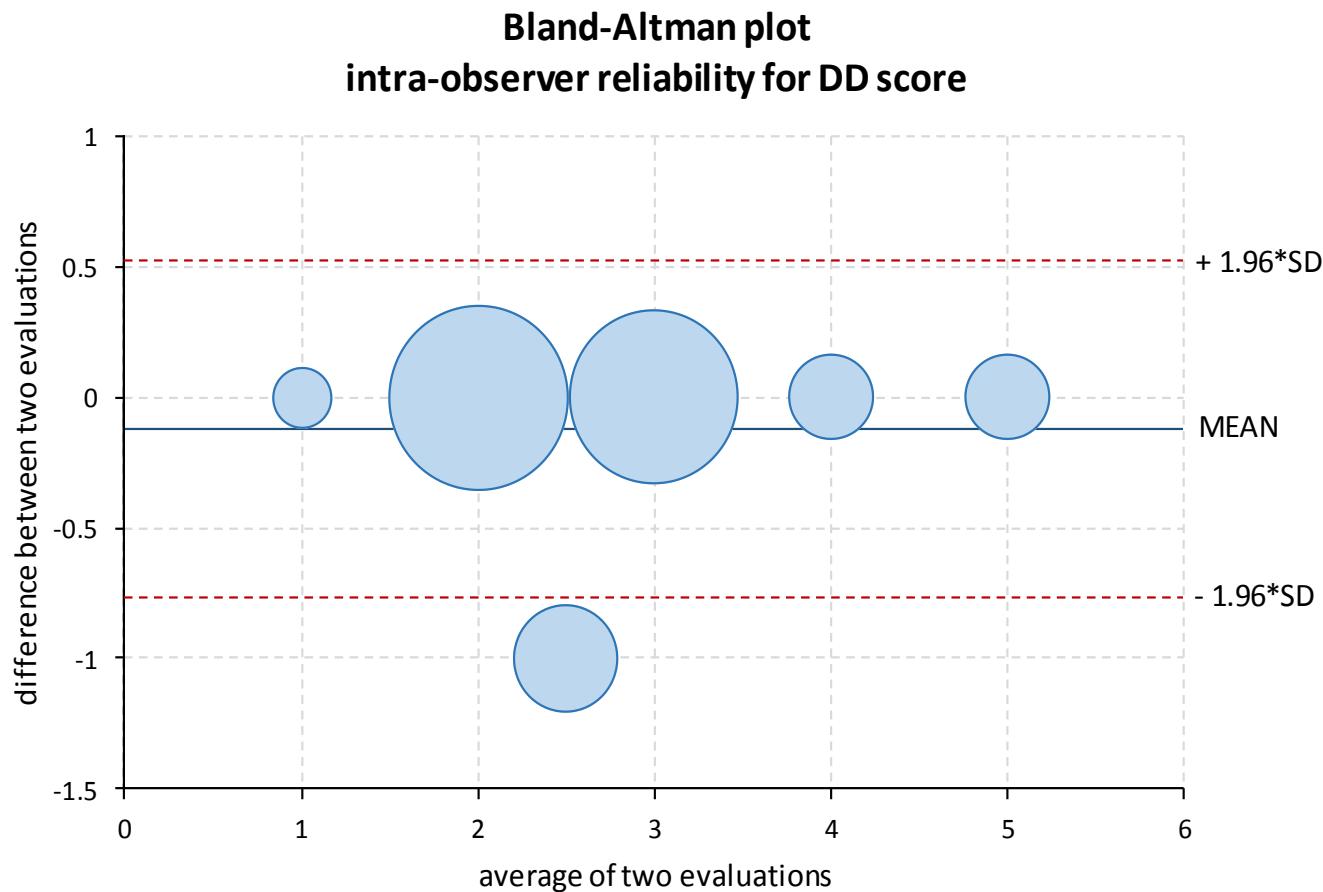
Supplementary Table 2 Analysis of inter-observer reliability

	Evaluation 1 ¹	Evaluation 2 ¹	Difference ¹	p-value ²	Correlation coefficient (p-value) ³
DD score	2.7 (\pm 1.0)	3.0 (\pm 0.7)	-0.28 (\pm 0.61)	0.035	0.673 (<0.001)
EP score	1.8 (\pm 1.7)	2.0 (\pm 1.5)	-0.16 (\pm 0.58)	0.059	0.732 (<0.001)
FD score	1.6 (\pm 0.9)	1.0 (\pm 0.7)	0.62 (\pm 0.60)	<0.001	0.771 (<0.001)
FS score	0.4 (\pm 0.7)	0.8 (\pm 0.9)	-0.40 (\pm 0.53)	<0.001	0.747 (<0.001)

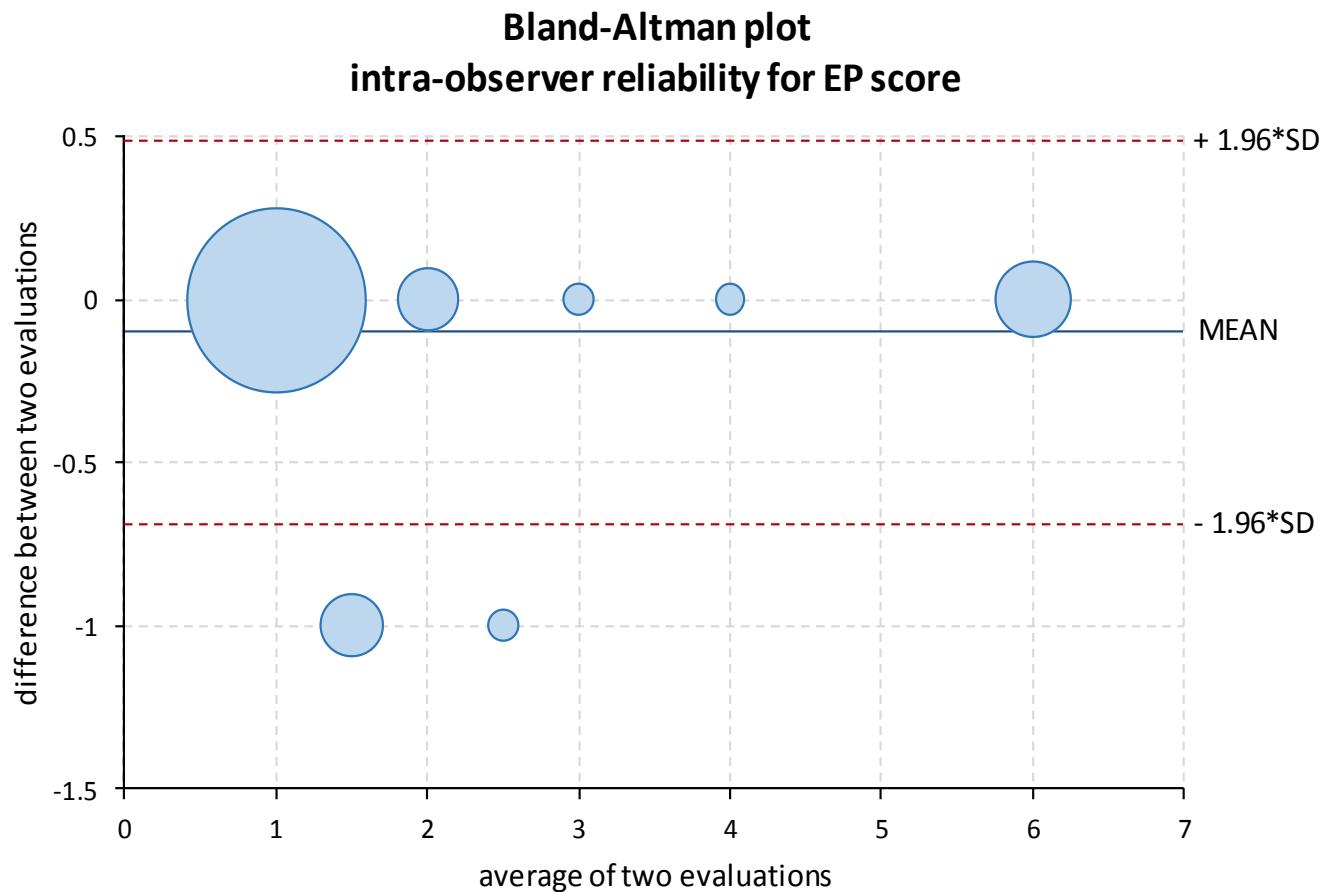
¹Mean (\pm SD)²Statistical significance of difference between observer 1 and observer 2, Wilcoxon signed-rank test³Spearman correlation coefficient

DD score – Disc degeneration score, EP score – Endplate score, FD score – Facet degeneration score, FS score – Foraminal stenosis score

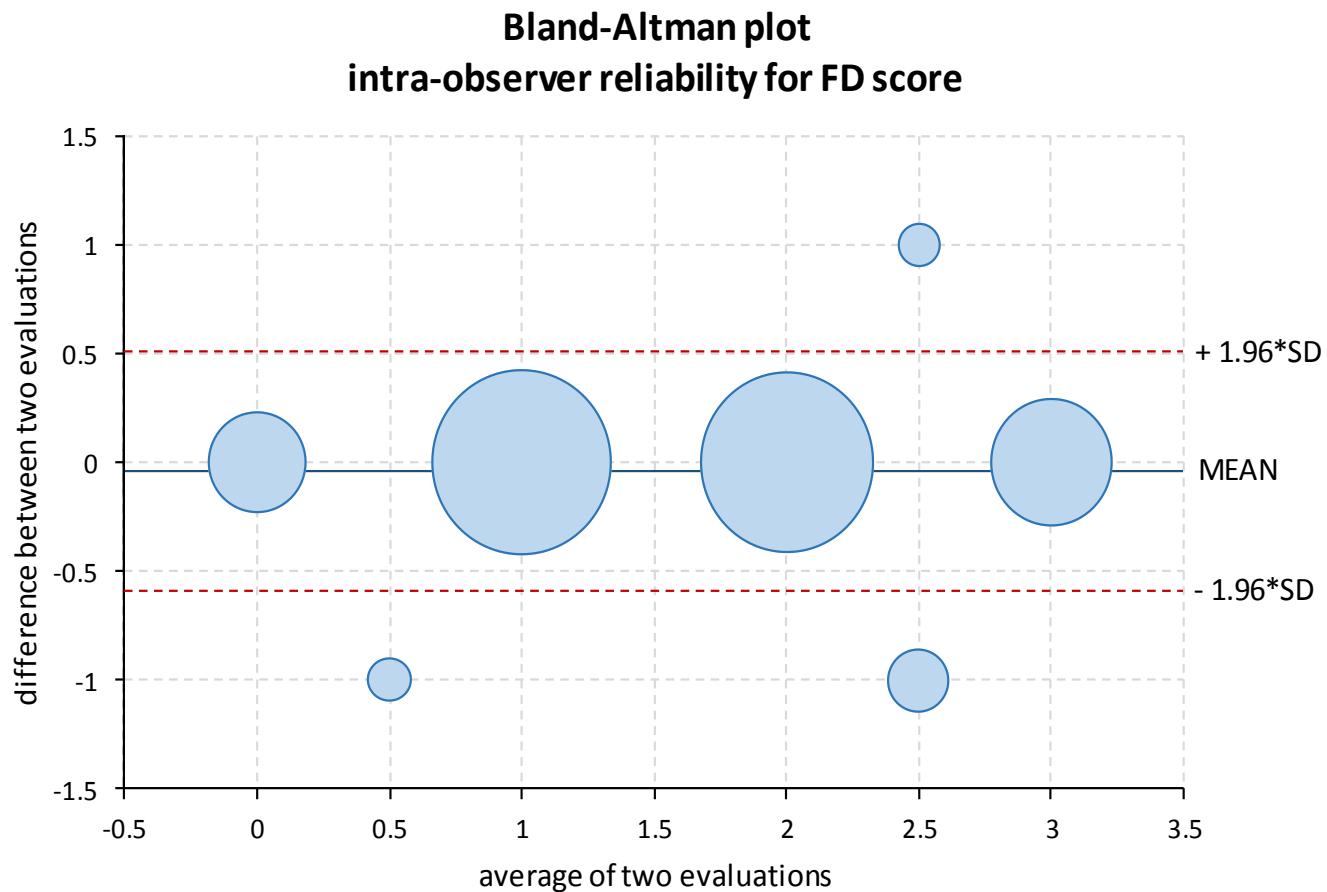
Supplemental Figure 1 Intra-observer reliability evaluated for DD score using Bland-Altman plots



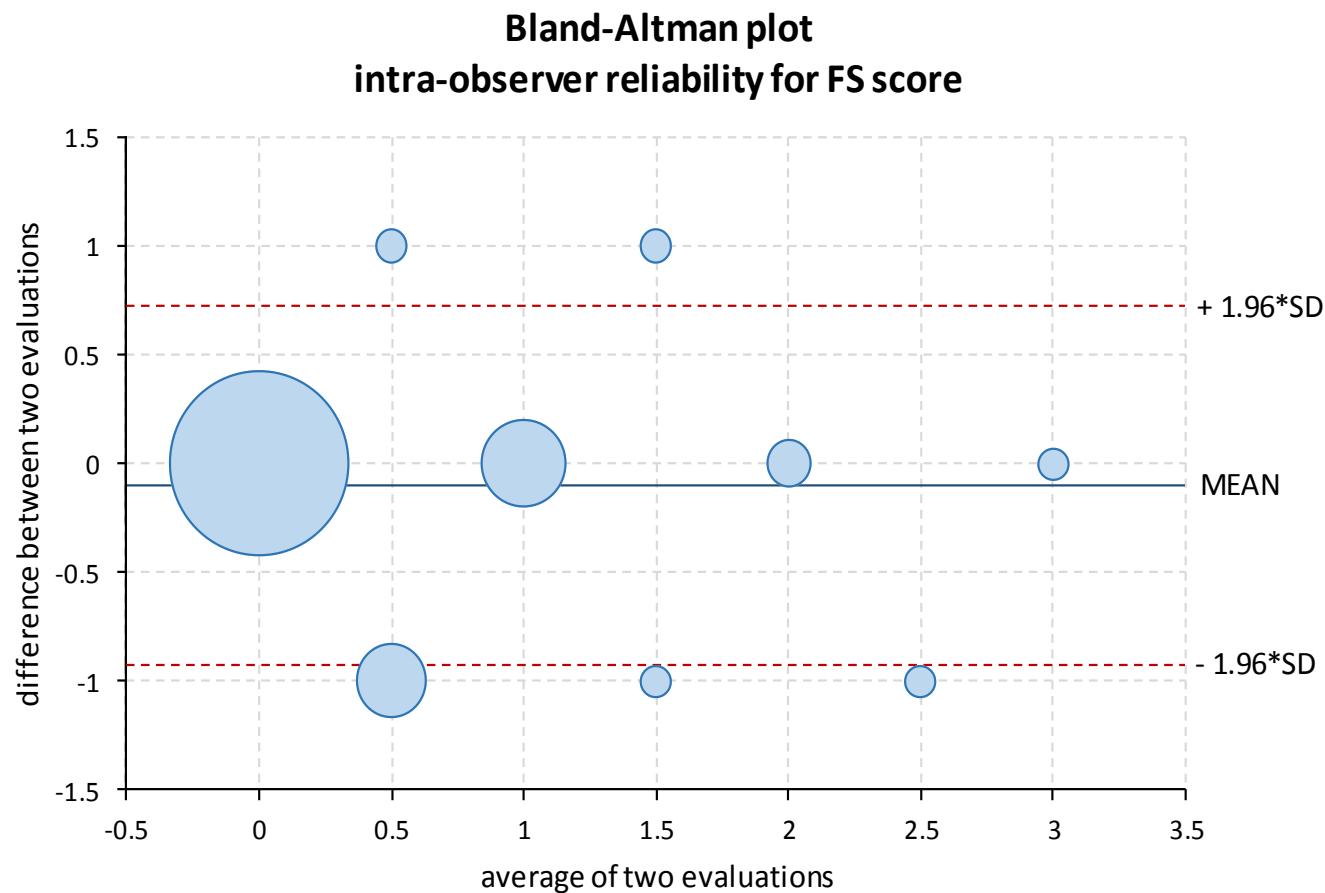
Supplemental Figure 2 Intra-observer reliability evaluated for EP score using Bland-Altman plots



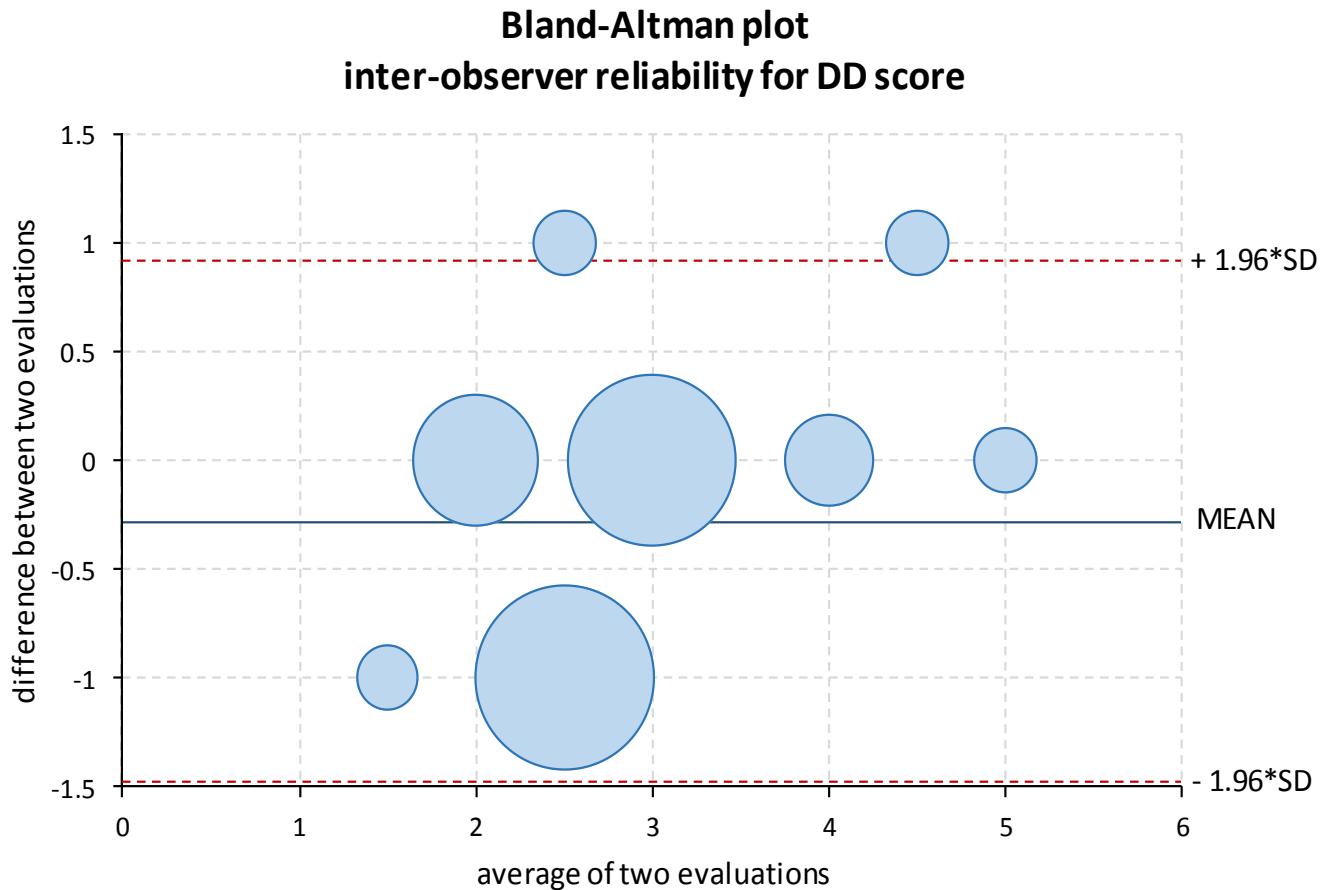
Supplemental Figure 3 Intra-observer reliability evaluated for FD score using Bland-Altman plots



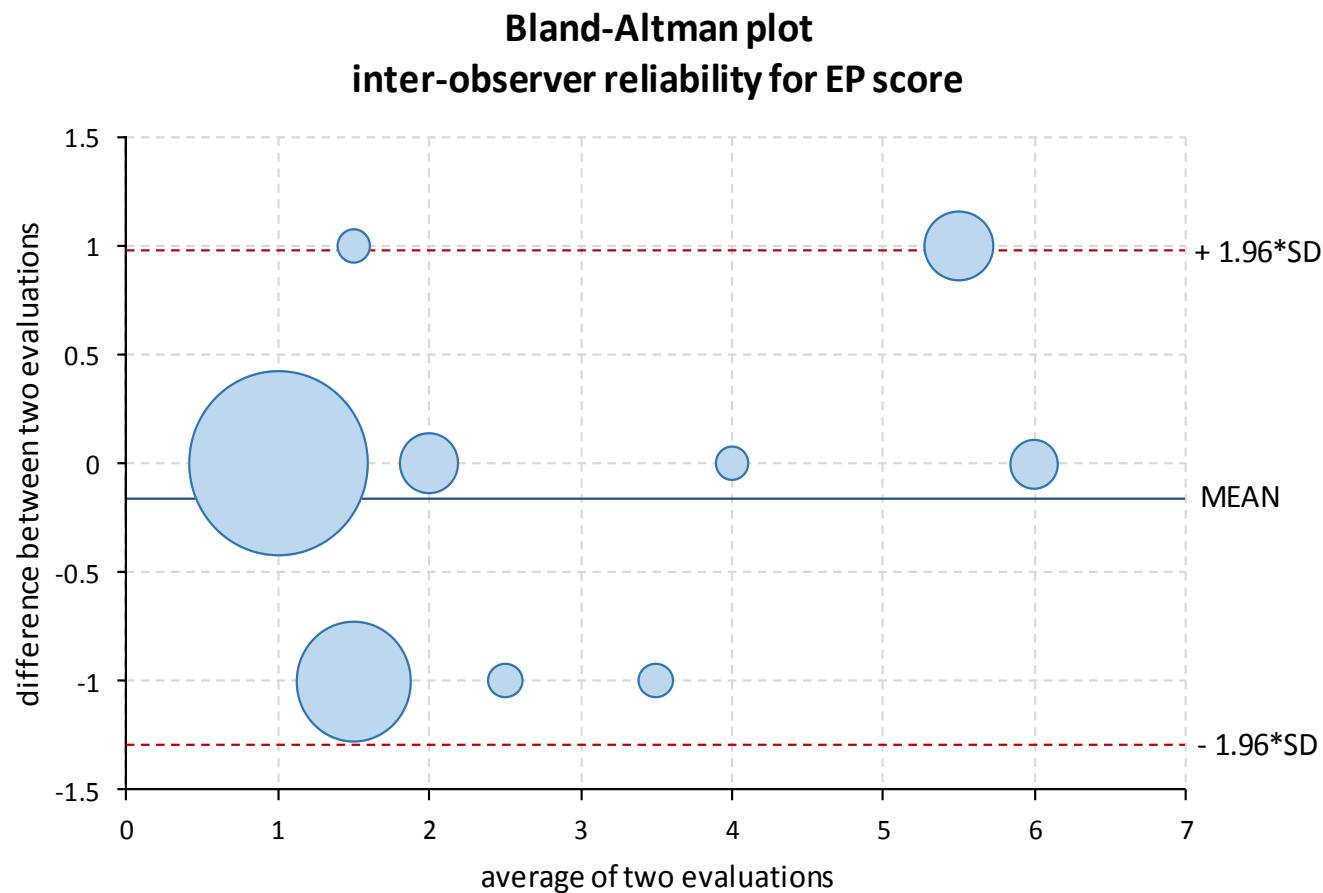
Supplemental Figure 4 Intra-observer reliability evaluated for FS score using Bland-Altman plots



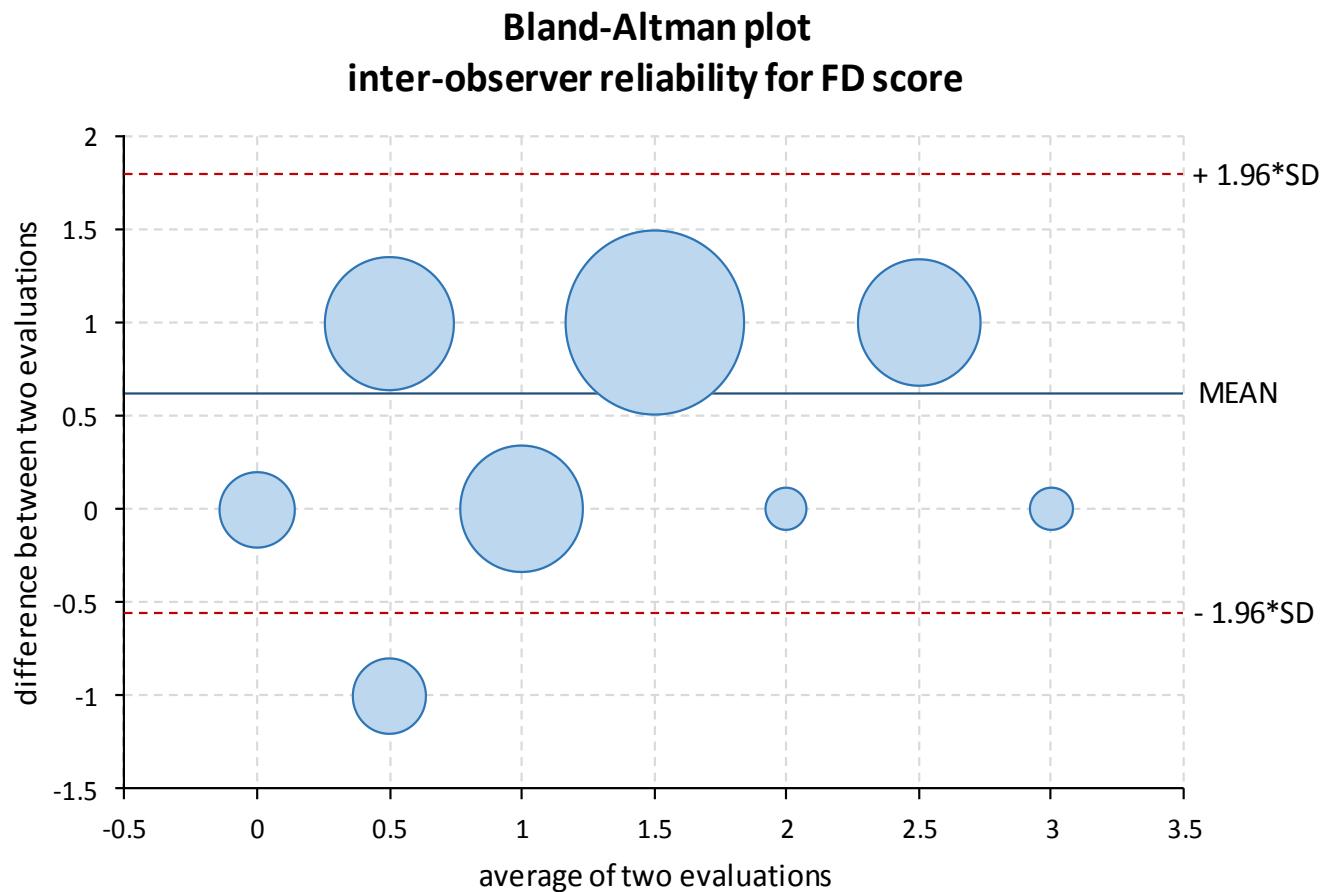
Supplemental Figure 5 Inter-observer reliability evaluated for DD score using Bland-Altman plots



Supplemental Figure 6 Inter-observer reliability evaluated for EP score using Bland-Altman plots



Supplemental Figure 7 Inter-observer reliability evaluated for FD score using Bland-Altman plots



Supplemental Figure 8 Inter-observer reliability evaluated for FS score using Bland-Altman plots

