

Appendix A Pairwise Comparisons of DSC Scores on Internal Test Data

Table A1: Tukey-adjusted pairwise comparisons of DSC scores between normalisation methods on internal test data, derived from a linear mixed-effects model. Positive estimates indicate higher DSC for the first-listed method.

Comparison	Estimated Δ DSC (%)	<i>p</i> -value
Z-score–CLAHE	0.214	0.030
Z-score–Robust min-max	0.146	0.331
Z-score–GMM	-0.007	1.000
Z-score–HE	0.361	< 0.001
Z-score–Nyúl	0.115	0.635
Z-score–Min-max	-0.006	1.000
CLAHE–Robust min-max	-0.068	0.957
CLAHE–GMM	-0.220	0.023
CLAHE–HE	0.147	0.330
CLAHE–Nyúl	-0.099	0.774
CLAHE–Min-max	-0.220	0.023
Robust min-max–GMM	-0.153	0.279
Robust min-max–HE	0.214	0.030
Robust min-max–Nyúl	-0.032	0.999
Robust min-max–Min-max	-0.153	0.279
GMM–HE	0.367	< 0.001
GMM–Nyúl	0.121	0.571
GMM–Min-max	0.000	1.000
HE–Nyúl	-0.246	0.006
HE–Min-max	-0.367	< 0.001
Nyúl–Min-max	-0.121	0.571