

Supplemental material

Cohen's κ with Cicchetti-Allison (CA) weighting was applied to evaluate the intra-observer agreement for each of the five observers and for image criteria 1-5. These kappa values were found to be rather poor (see Results) but clinical although acceptable. The reason for considering the results as clinical acceptable when the calculated agreement as assessed by kappa values was found to be rather poor is explained by a limitation of kappa values as the number of used categories influences the kappa value despite a constant proportion of concordant results. For instance, 2/3 of repeated readings were in concordance for reader 1, whereas 1/3 of repeated VGA scores were in disagreement for both criterion 1 and 4. The disagreeing results deviate by only one unit of the score for all images. In Table A1, all intra-observer assessments for reader 1 are shown. Reader 1 made use of 2 categories for criterion 1, but he used 4 categories for criterion 4. Despite the identical proportion of results in agreement, does the kappa value differ significantly due to its definition.

Table A1: Images repeated (intra-observer agreement results) for observer 1 image criteria 1 and 4 related to the calculated kappa value

| Patient | Reconstruction | Criteria 1 – 1st | Criteria 1 – 2nd | Criteria 4 – 1st | Criteria 4 – 2nd |
|--------------------|----------------|------------------|------------------|---|------------------|
| 1 | 30% ASIR | 4 | 5 | 5 | 5 |
| 4 | FBP | 4 | 4 | 4 | 4 |
| 4 | 30% ASIR | 4 | 4 | 3 | 4 |
| 7 | FBP | 4 | 4 | 2 | 2 |
| 7 | 30% ASIR | 5 | 4 | 5 | 5 |
| 8 | 60% ASIR | 5 | 5 | 5 | 5 |
| 12 | FBP | 4 | 4 | 3 | 4 |
| 23 | 60% ASIR | 4 | 5 | 4 | 4 |
| 26 | 60% ASIR | 4 | 4 | 2 | 3 |
| Kappa value | | 0.18 | | 0.55 (simple kappa); 0.72 (weighted kappa) | |