Purpose: Standard methods of computed tomography (CT) image-quality assessment in terms of objective performance parameters (e.g., noise, spatial-resolution, low contrast resolution, slice-thickness etc.) may lack in adequately incorporating the effect of patient shape, anatomical noise, and clinically relevant diagnostic task, such as detection of abnormal pathology. Most of the existing psychophysical experiments, e.g., 2-AFC, ROC, FROC, MRMC etc., are very expensive to conduct and subject to a combinatorial explosion, making routine human experiments impractical. In an effort to develop a realistic practical assessment of imaging-performance in clinically relevant diagnostic tasks, a practical human-observer-based scoring method was developed to quantify not only the probability of true lesion detection, but also an average detection score as an indicator of quality of detection in complex tasks, and applied in comparing two reconstruction technologies.

Methods: An anthropomorphic abdomen phantom with multiple simulated hyper-vascular liver-lesions of clinically relevant sizes and contrast-levels was used to evaluate human-observer-performance in a low-contrast-lesion-detection task. Human observer-based imaging performance assessment of the two reconstruction technologies, filtered-back-projection (FBP) and the VeoTM (model-based-iterative-reconstruction), at two dose levels, ”full dose” (corresponding to dose levels in routine clinical practice for abdominal CT) and 1/4th of “full dose”, were carried out. Average lesion detection performance in a task with multiple lesions of different contrast-levels and sizes, quantified in terms of average-probability-of-detection and average-detection-score as figures-of-merit, was evaluated for relative imaging performance comparison with-respect-to dose.

Results: Results of the study and statistical significance tests indicated that average-probability-of-detection and average-detection-score at full dose FBP and 1/4th dose VeoTM are comparable, whereas at the same full dose VeoTM offers improved detection performance.

Conclusions: The presented human-observer-based scoring method has significant potential for practical imaging-performance comparison of different reconstruction technologies under clinically relevant tasks, where single image may have multiple pathologies of different contrast-levels, shapes and sizes.