

Title: Evaluating reliability of structured computed tomography scan reporting in pancreatic cancer staging: insights from a quality assurance study exploring inter- and intra-observer agreement and variability

Ellis SJ^{1,2}, Alam F³, Kasza J³, Li L³, Di Muzio B¹, Gerstenmaier J¹, Goodwin M⁴, Ko HS⁵, Lim K⁴, Ioannou L³, Recasens A³, Zalcborg J^{3,6}, Pilgrim CHC^{2,3,7}

1. Department of Radiology, Alfred Health, Melbourne, Victoria, Australia
2. Department of Surgery, School of Translational Medicine, Monash University
3. School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
4. Department of Radiology, Austin Hospital, Australia
5. Department of Cancer Imaging, The Peter MacCallum Cancer Centre, Melbourne, VIC, Australia
6. Department of Medical Oncology, Alfred Health, Melbourne, Victoria, Australia
- Introduction
7. Department of Surgery, School of Translational Medicine, Monash University

Introduction: The purpose of this study is to evaluate the inter-observer and intra-observer agreement of a standardised synoptic report that follows the 2017 International Consensus definition for borderline resectable pancreatic ductal adenocarcinoma (PDAC).

Methods: Sixty dedicated pancreatic protocol CT scans were independently reported by six radiologists using the standardised synoptic report. Variability was indicated by percentage agreement between anatomical findings and final resectability status. Gwet's Agreement Coefficients (AC) was calculated to determine inter-observer agreement and Cohen's Kappa for intra-observer agreement.

Results: A total of 22320 scan parameters were collected, with six radiologists individually rating 60 scans. Twenty-six scans had previously been rated in the pilot study by four radiologists and were re-rated with radiologists blinded to their previous response. Full inter-observer agreement of anatomical resectability classification was seen in 21.67% (13/60) scans, with 78.33% (47/60) showing majority inter-observer agreement. The intra-observer agreement of anatomical resectability classification was 65.38% (17/26). The highest level (very good) of agreement was seen regarding biliary stent presence (Gwet's AC 0.96, 95% CI 0.93, 1.00) and the least (poor) significant agreement was seen describing jejunal vein involvement (Gwet's AC 0.17, 95% CI 0.00, 0.35). Moderate inter-observer and intra-observer agreement for overall resectability status was demonstrated with Gwet's AC 0.40 (95% CI 0.29, 0.51) and Cohen's Kappa 0.45 (95% CI 0.20, 0.69).

Conclusion: This study highlights synoptic reporting variability pertaining to PDAC resectability status, with moderate inter- and intra-observer agreement elucidating inherent complexities, particularly regarding jejunal vein assessment.