A Quantitative Metric to Predict Patient-Specific QA Failure in Lung SBRT Integrated into the Planning Process

Caroline Colbert, PhD - University of Washington; Eric Ford, PhD - University of Washington; Minsun Kim, PhD - University of Washington

Objectives: Stringent patient-specific quality assurance (PSQA) is an essential component of a comprehensive SBRT QA program. However, features unique to SBRT treatments, including small field sizes and highly modulated treatment plans, create challenges to reliable 3D array measurement-based PSQA. Failed PSQA measurements necessitate re-planning and repeated PSQA, straining clinical resources and potentially delaying the start of treatment. A quantitative metric to reliably identify lung SBRT QA plans with a high likelihood of failure may allow for upfront creation of alternative SBRT treatment plans, preventing treatment delays for patients. This work evaluates a quantitative QA failure prediction metric that does not introduce additional overhead to the process time of routine clinical practice.

Methods: We identified 27 consecutive clinical SBRT-VMAT lung treatment plans (n=54 beams) completed between 2022 and 2023 in our clinic. All treatment plans were created by certified medical dosimetrists for delivery on a clinical C-arm LINAC. PSQA was performed and analyzed with a gamma criterion of 3%, 2mm and a dose threshold of 10%, in accordance with the recommendations of AAPM Task Group 218. We evaluated the mean leaf gap (MLG) for the smallest 10%, 25%, and 50% of all open leaf pair gaps (MLG-S10, MLG-S25, MLG-S50), as well as the MLG for the largest 10%, 25%, and 50% of all open leaf pair gaps (MLG-L10, MLG-L25, MLG-L50). These metrics compute the MLG of a subset of leaf pair gaps in all VMAT segments, where leaf gap is defined as the distance between unobscured opposing leaf pairs. The unweighted mean of all leaf gaps meeting specific criteria was calculated using a python-based scripting protocol within the treatment planning system. We also evaluated PTV volume as a predictor of PSQA failure.

Results: We found a mean gamma pass rate of 84.68±7.42%. Thirteen plans (48%) failed QA by institutional criteria (gamma pass rate < 85%). ROC analysis of all MLG metrics showed AUC ranging from 0.533 (95% CI 0.302, 0.764) with MLG-S10 to 0.857 (95% CI 0.708, 1.00) with MLG-L25. Using a criterion of MLG-L25 < 2.93 cm, we find a sensitivity of 0.846 and a specificity of 0.714 Regression analysis showed a significant correlation between this MLG metric and gamma pass rate (R^2=0.49, $p < 0.001$). PTV volume ranged from 4.5 to 58.9 cc. The distribution of PTV volume was skewed small, with 21 PTVs (78%) having a volume less than 20 cc. While correlated with gamma pass rate (R^2=0.33, $p < 0.05$), PTV volume proved a poorer predictor of plan QA failure than the MLG-L leaf gap metrics (AUC 0.742, 95% CI 0.550, 0.994). All lung SBRT-VMAT plans with a PTV volume greater than 20 cc (N=6) passed PSQA. However, after excluding these plans PTV volume was not significantly correlated with gamma pass rate (R^2=0.04, p=0.37).

Conclusion(s): Within the range of PTV volumes typical of lung SBRT treatments in our clinic, PTV volume is not predictive of PSQA gamma pass rate. The MLG metrics that capture the upper range, rather than the lower range, of MLC leaf gaps are better predictive of 3D array measurement-based PSQA failure. MLG-L75 is a simple quantitative metric that shows promise as a predictor of PSQA failure in lung SBRT-VMAT treatment plans. Evaluating this plan metric in the treatment planning scripting environment requires minimal additional effort. Accurately identifying plans likely to fail PSQA has the potential to provide for upfront replanning and workload savings.