Author Disclosure: E.M. Chung: None. A. Demurchyan: None. D.J. Lu: None. A.T. Nguyen: None. E. Anderson: None. K.M. Atkins: None. M. Kamrava: American Board of Radiology, Brachytherapy Journal, American Brachytherapy Society Board of Directors. Co-Editor; Journal of Contemporary Brachytherapy.

129

Clinical Validation of Deep Learning Algorithms for Lung Cancer Radiotherapy Targeting

A. Hosny,[1] D.S. Bitterman,[2] C.V. Guthier,[3] H. Roberts,[4] S. Perni,[4] A. Saraf,[5] J.M. Qian,[6] L.C. Peng,[7] I.M. Pashtan,[8] B.H. Kann,[1] D. Kozono,[9] P. Catalano,[4] H. Aerts,[10] and R.H. Mak[3]; *[1]Dana-Farber Cancer Institute, Boston, MA, [2]Harvard Radiation Oncology Program, Boston, MA, [3]Department of Radiation Oncology, Brigham and Women's Hospital and Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, [4]Dana Farber Cancer Institute, Boston, MA, [5]Massachusetts General Hospital, Boston, MA, [6]BWH/DFCI/MGH - Harvard Radiation Oncology Program, Boston, MA, [7]Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, [8]Brigham and Women's Hospital and Dana-Farber Cancer Institute, Boston, MA, [9]Brigham And Women's Hospital, Boston, MA, [10]Cardiovascular Imaging Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA*

Purpose/Objective(s): Automated target segmentation for non-small cell lung cancer (NSCLC) patients has the potential to support radiation treatment planning. Artificial intelligence (AI) has demonstrated great promise in medical image segmentation tasks. However, most studies have been confined to *in silico* validation in small internal cohorts, lacking data on real-world clinical utility. In this study, we developed primary tumor and involved lymph node segmentation algorithms in computed tomography (CT) images. Validation is performed in multiple large multi-institutional cohorts to assess model generalizability.

Materials/Methods: Simulation CTs and ground truth annotations were collected from multiple public and private sources (total n = 2584). We employed the following benchmarks: Inter-observer (6 radiation oncologists, n = 20, median volumetric dice 0.83, 95% CI 0.82-0.84) and intra-observer (1 radiation oncologist, 3 reads, n = 21, median volumetric dice 0.88, 95% CI 0.84-0.9). We developed two segmentation algorithms: seed-point assisted and fully automated. Model training data (n = 787) comprised NSCLC-Radiomics (stages I-IIIB, n = 422) and LungRT-1 (stages IA-IV, n = 365). Validation was first performed in an internal dataset annotated by a single thoracic radiation oncologist (LungRT-1, n = 136). Additional validation included: (1) an internal dataset annotated by other radiation oncologists, including generalists, in our center (LungRT-2, n = 1075), (2) an external clinical trial dataset from 185 different institutions (RTOG-0617, n = 403), and (3) a dataset of early-stage surgical patients annotated for diagnostic purposes by radiologists (NSCLC-Radiogenomics, n = 142). Volumetric dice, using expert manual segmentations as ground truth, was used as an evaluation metric.

Results: The model performance is comparable to the benchmarks when validated on internal data, with degrading performance in cohorts annotated by other radiation oncologists.

Conclusion: The results highlight the importance of assessing segmentation style among annotators and understanding model generalizability in external cohorts, all while cautioning against degrading performance in increasingly external data. Differences between radiologists and radiation oncologists performing the same segmentation task underscore the importance of clinical context in AI model deployment. Further validation includes studying the dosimetric impact of AI-generated segmentations, and conducting human subject experiments to assess AI output acceptance and time savings.

Abstract 129 — Table 1

Dataset	Stage (I, II, III, IV, n/a %)	Seed Point Assisted Dice	Fully Automated Dice	*P*-value (seed point assisted dice vs inter-observer benchmark)
LungRT-1	23, 5, 60, 10, 2	0.83 (0.82-0.85)	0.82 (0.80-0.83)	0.9
LungRT-2	12, 8, 46, 32, 2	0.61 (0.59-0.63)	0.59 (0.57-0.61)	< 0.001
RTOG-0617	0, 0, 93, 0, 7	0.71 (0.69-0.73)	0.69 (0.67-0.72)	< 0.001
NSCLC-Radiogenomics	34, 27, 10, 4, 25	0.68 (0.63-0.73)	0.64 (0.59-0.69)	< 0.001

Author Disclosure: A. Hosny: Consultant; Altis Labs. Stock; Altis Labs. D.S. Bitterman: Honoraria; Agios Pharmaceuticals. C.V. Guthier: None. H. Roberts: None. S. Perni: None. A. Saraf: None. J.M. Qian: None. L.C. Peng: None. I.M. Pashtan: None. B.H. Kann: None. D. Kozono: None. P. Catalano: None. H. Aerts: Stock; Onc. Ai. R.H. Mak: Research Grant; ViewRay, Inc. Honoraria; NewRT, ViewRay, Inc. Advisory Board; AstraZeneca. Travel Expenses; NewRT.

130

Final Results of a Phase I "RadVax" Trial of Hypofractionated Radiation Combined With Pembrolizumab in Patients With Metastatic Solid Tumors

J.N. Lukens,[1] R. Mick,[2] A.C. Huang,[3] N. Han,[4] M.D. Farwell,[5] T.C. Mitchell,[6] R. Amaravadi,[7] L.M. Schuchter,[3] A.T. Berman,[8] M. O'Hara,[9] A. Maity,[10] D. Miller,[8] A.J. Minn, III[7] R.H. Vonderheide,[6] E.J. Wherry,[11] and A. Maity[8]; *[1]University of Pennsylvania, Department of Radiation Oncology, Philadelphia, PA, [2]Department of Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA, [3]Department of Medicine, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA, [4]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, [5]Department of Radiology, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA, [6]Department of Medicine, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA, [7]University of Pennsylvania, Philadelphia, PA, [8]Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA, [9]Department of Hematology and Oncology, University of Pennsylvania, Philadelphia, PA, [10]Department of Radiation Oncology, Sidney Kimmel Medical College & Cancer Center at Thomas Jefferson University, Philadelphia, PA, [11]Department of Microbiology, Perelman School of Medicine of the University of Pennsylvania, Philadelphia, PA*

Purpose/Objective(s): Many patients treated with anti-PD-1 therapy do not show a clinical response. Preclinical studies suggest that adding hypofractionated radiotherapy (HFRT) to anti-PD1 can increase the efficacy of immunotherapy through several mechanisms including increased antigen presentation. We conducted a prospective trial testing the combination of pembrolizumab and HFRT in patients with metastatic solid tumors.

Materials/Methods: This prospective single-institution phase I trial tested pembrolizumab in combination with HFRT in patients with metastatic cancers (NSCLC, melanoma, pancreas, breast, others) and an ECOG performance status of 0-1. Melanoma and NSCLC patients were required to have progression of disease on anti-PD1, having received ≥ 2 doses of anti-PD1 and progression documented by RECIST v1.1. Patients were required to have an index lesion ≥1 cm that was amenable to HFRT and at least one other lesion that was not irradiated and could be followed for response using RECIST criteria. Pembrolizumab 200 mg IV every 3 weeks was administered beginning 1 week prior to the first fraction of radiation. The HFRT dose was 8 Gy x 3 fractions or 17 Gy x 1 fraction, determined by randomization during the Expansion phase. The primary objective was the safety of HFRT combined with pembrolizumab, with dose-limiting toxicity (DLT) defined as Grade ≥ 3 non-hematological toxicity related to