

1300

CT SCAN NORMALIZATION TO IMPROVE RADIOMICS GENERALIZABILITY AND ASSESS PROGRESSION-FREE SURVIVAL IN PATIENTS WITH ADVANCED NON-SMALL CELL LUNG CANCER TREATED WITH IMMUNE CHECKPOINT INHIBITORS: A MULTICENTRIC STUDY

¹Marion Tonneau*, ²Kim Phan, ²Cécile Low-Kam, ²Francis Dutil, ³Suzanne Kazandjian, ⁴Davy Vanderweyden, ³Justin Panasci, ¹Julie Malo, ⁵François Coulombe, ¹Arielle Elkrief, ¹Wiam Belkaïd, ²Lisa Di Jorio, ⁵Michèle Orain, ⁶Nicole Bouchard, ³Thierry Muanza, ²Kam Kafi, ²Florent Chandelier, ³Philippe Joubert, ¹Bertrand Routy. ¹CRCHUM, Montréal, Canada; ²Imagia Canexia Health, Montreal, Canada; ³Jewish General Hospital, Montreal, Canada; ⁴Centre Hospitalier de Sherbrooke, Sherbrooke, Canada; ⁵IUCPQ, Québec, Canada; ⁶CHUS, Sherbrooke, Canada

Background Recent development in the field of artificial intelligence suggests that radiomics may represent a promising non-invasive biomarker to predict response to immune checkpoint inhibitors (ICI). Nevertheless, validation of radiomics algorithms in independent cohorts remains a challenge in part due to variations in image acquisition and reconstruction. Using radiomics as a biomarker for ICI response in non-small cell lung cancer (NSCLC) patients, we investigated the importance of scan normalization as part of a broader statistical framework to enable model external generalizability.

Methods Discovery cohort was composed of pre-ICI CT scans of 514 advanced NSCLC patients from three academic centers (n=223 CHUM, n=130 JGH and n=161 IUCPQ). A validation cohort included 144 patients from a fourth center (CHUS). Power calculation for the validation cohort was determined with an AUC threshold of 0.67 in the discovery cohort to predict progression-free survival at 6 months (PFS-6), alpha-risk was 0.10, leading to n=140 patients. Radiomics features were extracted from original clinical scans using the established open source PyRadiomics, and a proprietary DeepRadiomics technology (leveraging deep learning). We harmonized images to account for variations in reconstruction kernels, slice thicknesses, and device manufacturers. Multivariable models, evaluated using leave-one-center-out cross validation, were used to estimate the predictive value of clinical variables (ECOG, line of treatment, stage, smoking), PD-L1 expression, and PyRadiomics or DeepRadiomics for PFS-6.

Results In this cohort, the best prognostic factor for PFS-6 excluding radiomics, was combination of clinical + PD-L1 expression with an AUC of 0.66 and 0.62 in the discovery and validation cohort. Without image harmonization, combining clinical + PyRadiomics or DeepRadiomics, AUC were 0.69 and 0.69 respectively in the discovery cohort, but depicting significant drops to 0.57 and 0.52, respectively in the validation cohort. This lack of generalizability was consistent with observations in principal component analysis clustering by center. Subsequently, image harmonization eliminated these clusters and when combining clinical + DeepRadiomics model, we obtained an AUC of 0.67 and 0.63 in the discovery and validation cohort, respectively. The combination of clinical + PyRadiomics failed generalizability validations, with AUC of 0.66 and 0.59, without improvement with the addition of PD-L1 (table 1).

Conclusions We demonstrated that a risk prediction model that combined clinical + DeepRadiomics was generalizable and had similar performances to clinical + PD-L1, when scan harmonization methods were in place. Altogether, this study showed the strong potential of radiomics as a future non-invasive strategy to predict ICI response in advanced NSCLC.

Ethics Approval Number of the approval: MP-02-2019-8091

Abstract 1300 Table 1 Area under the curve (AUC) and confident interval of clinical, clinical + PD-L1, clinical + PyRadiomics and clinical + DeepRadiomics models, in the discovery and validation cohorts, before and after scans harmonization.

Table 1 – Area under the curve (AUC) and confident interval of clinical, clinical + PD-L1, clinical + PyRadiomics and clinical + DeepRadiomics models, in the discovery and validation cohorts, before and after scans harmonization.

	Before scans harmonization		After scans harmonization	
	Discovery cohort	Validation cohort	Discovery cohort	Validation cohort
Clinical	0.64 95%CI [0.59-0.69]	0.58 95%CI [0.49-0.69]	-	-
Clinical + PD-L1	0.66 95%CI [0.61-0.70]	0.62 95%CI [0.51-0.71]	-	-
Clinical + PyRadiomics	0.69 95%CI [0.64-0.74]	0.57 95%CI [0.48-0.67]	0.66 95%CI [0.61-0.70]	0.59 95%CI [0.49-0.68]
Clinical + DeepRadiomics	0.69 95%CI [0.65-0.74]	0.52 95%CI [0.43-0.62]	0.67 95%CI [0.63-0.72]	0.63 95%CI [0.53-0.73]

<http://dx.doi.org/10.1136/jitc-2022-SITC2022.1300>