Clinical integration and validation of an AI-based radiomic platform for predicting PD-(L)1 immune checkpoint inhibitor response in stage IV NSCLC

Chiharu Sako, Ph.D. [1], Lester Boeh, M.B.A. [1], Chong Duan, Ph.D. [2], Kevin Maresca, Ph.D. [2], Subha Madhavan, Ph.D. [2], Ray D. Page, D.O., Ph.D. [3], Allen Ardestani, M.D., Ph.D. [4], Ravi B. Parikh, M.D., M.P.P. [5], George R. Simon M.D. [6], Hugo J.W.L. Aerts, Ph.D. [1,7], Petr Jordan, Ph.D. [1],
[1]Onc.AI, San Carlos, CA, USA; [2]Pfizer, Cambridge, MA, USA; [3]The Center for Cancer and Blood Disorders, Weatherford, TX, USA; [4]West Cancer Center and Research Institute, Germantown, TN, USA; [5]University of Pennsylvania, Philadelphia, PA, USA; [6]Moffitt Cancer Center, Tampa, FL, USA; [7]Mass General Brigham, Harvard Medical School, Boston, MA, USA

Introduction

There is an urgent clinical need to identify patients likely to benefit from immune checkpoint inhibitor (ICI) treatment. Approaches available in the clinic today, such as PD-L1 immunohistochemistry (IHC) and tumor mutation burden (TMB), are insufficient for this task, in part as differences in microenvironments expressed by individual tumors may lead to heterogeneous response patterns. Recent efforts exploring the utility of quantitative imaging (radiomic) biomarkers to predict response to ICIs have shown promise to provide a more accurate and scalable method.

Methods

Development of clinical decision support platform. We have developed an artificial intelligence (AI)-based clinical decision support platform intended to be used by medical oncologists for planning and guidance of ICI treatment plans in first-line stage IV NSCLC patients. The platform is a digital-only, cloud-hosted software-as-a-service, which medical oncology clinics can access entirely over the internet (see **Error! Reference source not found.**). We have developed predictive models for lesion-level [1] and patient-level [2] ICI therapy response using a large multi-institutional retrospective dataset [3] and are currently validating technical performance on independent real-world datasets, retrospective clinical trial datasets, as well as undergoing clinical integration testing and optimization.

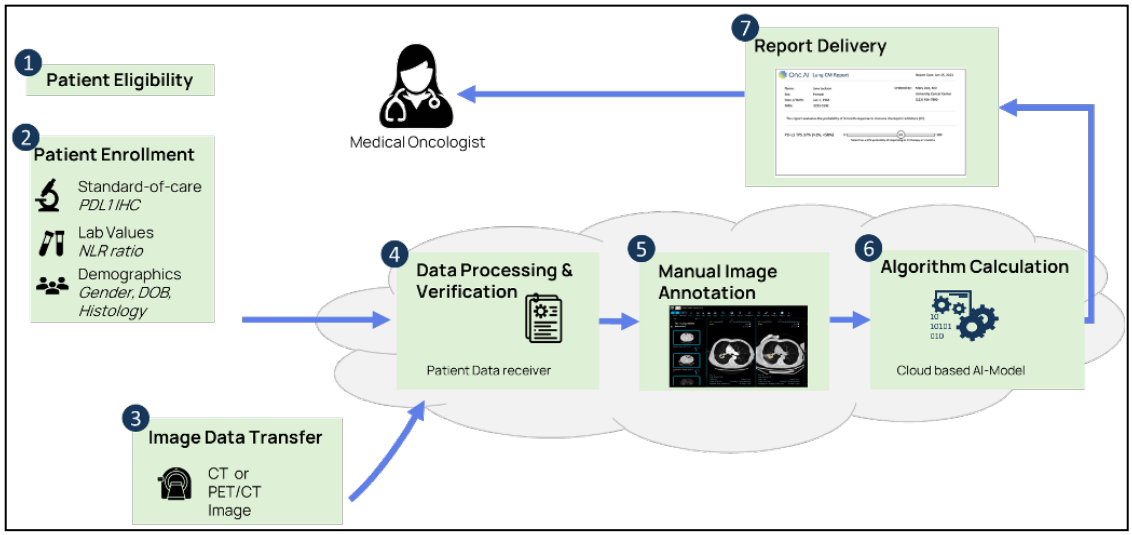

Figure 1. Clinical integration of an AI-based radiomic platform for planning and guidance of ICI treatment plans.

Technical performance on multi-institutional real-world datasets. Using a large multi-institutional real-world dataset, we analyzed radiomic characteristics of 6,295 primary and metastatic lesions from 1,206 stage IV NSCLC patients treated with PD-(L)1 ICIs from nine institutions, representing more than 50 clinics across the US and Europe. The data aggregation protocol was designed to capture a broad range of demographic, geographic, and imaging protocol variables representative of the full spectrum of clinical treatment patterns ranging from large academic institutions to community cancer centers. Patients with unavailable imaging follow-up or with oncogenic driver mutations were excluded from analysis, resulting in a total dataset of 791 subjects randomly assigned to training (N=541) and validation sets (N=250). Radiomic data was extracted from baseline CT scans capturing tumor heterogeneity, spicularity, and burden in the lung, lymph nodes, and liver. A multi-modal ensemble classifier

combining demographic features, PD-L1 TPS, and radiomic data was developed to predict response to ICI therapy per RECIST 1.1 criteria. The model's performance was evaluated in terms of the area under the receiver operating characteristic curve (ROC-AUC) and compared to PD-L1 IHC using the two-tailed DeLong test.

Validation on retrospective clinical trial data

The models developed on the large multi-institutional real-world dataset are being evaluated using an independent cohort from a completed clinical trial (NCT02573259) to explore the model generalizability to novel PD-(L)1 ICI agents, such as Sasanlimab.

Early phase of integration within clinical workflows and delivery of insights to clinicians

The primary goal of developing the platform is to take the next step in personalized care by supporting clinicians in the treatment decision making by providing additional insights that are incremental to the currently available data. Retrospective pilot studies have been initiated at two clinical institutions to evaluate the optimal integration into the clinical workflow in different treatment settings and deliver insights to clinicians when and where they are needed.

Results

Under the two-tailed DeLong test, the multi-modal model demonstrated statistically significant benefit over the current standard of care (PD-L1 IHC) in predicting multi-lesion 3-month response (Table 1): 0.81 (P=.005) area under the receiver operating characteristic curve (AUC) in first-line ICI monotherapy patients, 0.72 (P=.044) in all-lines ICI monotherapy, and 0.71 (P=.025) in all-lines ICI-chemotherapy combination. The imaging-only model demonstrated predictive performance comparable to PD-L1 IHC: 0.71 (P=.226), 0.61 (P=.905), 0.62 (P=.674) on the same cohorts respectively.

Table 1. Biomarker predictive performance of 3-month multi-lesion response per RECIST 1.1.

Biomarker	First-line ICI monotherapy, N=91 AUC (95% CI)	All-lines ICI monotherapy, N=138 AUC (95% CI)	All-lines ICI plus chemotherapy, N=114 AUC (95% CI)
Multi-modal	0.81** (0.69-0.92)	0.72* (0.62-0.82)	0.71* (0.60-0.83)
CT radiomics	0.71 (0.58-0.84)	0.61 (0.49-0.72)	0.62 (0.47-0.76)
PD-L1 IHC	0.60 (0.48-0.73)	0.60 (0.49-0.70)	0.58 (0.50-0.66)

*/** indicates statistical significance of comparison to PD-L1 IHC at the 5%/1% level under two-tailed DeLong test.

Conclusion

A multi-modal CT radiomics-based AI approach demonstrated predictive accuracy benefit over the current clinical standard and may provide an opportunity for more personalized patient management, such as risk-based escalation/de-escalation of concurrent chemotherapy in NSCLC patients.

The evidence generated using retrospective real-world datasets in the first phase of development was sufficiently compelling to motivate investigation of the predictive multi-modal models on novel ICI agents and combinations. Predictive performance evaluation using retrospective Sasanlimab clinical trial dataset is currently ongoing and will be reported using similar patient-level and lesion-level methodology.

References

1. Parikh R, Jordan P, Ciaravino R, et al. Radiomics-based multi-modal prediction of treatment response to PD-1/PD-L1 immune checkpoint inhibitor (ICI) therapy in stage IV non-small cell lung carcinoma (mNSCLC). Journal for ImmunoTherapy of Cancer 2022;10.
2. Parikh R, Jordan P, Ciaravino R, et al. Multi-institutional validation of a radiomics-based artificial intelligence method for predicting response to PD-1/PD-L1 immune checkpoint inhibitor (ICI) therapy in stage IV NSCLC. AACR Annual Meeting, Orlando, FL, 2023.
3. Sako C, Jordan P, McCall R, et al. Multi-center real-world data curation and assessment of tumor growth rate and overall survival in advanced NSCLC treated with PD-(L)1 immune checkpoint inhibitor therapy. ASCO Annual Meeting, Chicago, IL, 2023. (submitted)