PATHOLOGICA 2024;116:222-231;
DOI: 10.32074/1591-951X-998

Original article

A preliminary study on the diagnostic performance of the uPath PD-L1 (SP263) artificial intelligence (AI) algorithm in patients with NSCLC treated with PD-1/ PD-L1 checkpoint blockade

Alessio Cortellini[1,2,3], Claudia Zampacorta[4,5], Michele De Tursi[6], Lucia R. Grillo[7], Serena Ricciardi[8], Emilio Bria[9,10], Maurizio Martini[11], Raffaele Giusti[12], Marco Filetti[13,14], Antonella Dal Mas[15], Marco Russano[1], Filippo Gustavo Dall'Olio[16], Fiamma Buttitta[4,5,17], Antonio Marchetti[4,5,17]

[1] Operative Research Unit of Medical Oncology, Fondazione Policlinico Universitario Campus Bio-Medico, Roma, Italy; [2] Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Roma, Italy; [3] Department of Surgery and Cancer, Hammersmith Hospital Campus, Imperial College London, London, United Kingdom; [4] Center for Advanced Studies and Technology (CAST), University Chieti-Pescara, Italy; [5] Diagnostic Molecular Pathology, Unit of Anatomic Pathology, SS Annunziata Hospital, Chieti, Italy; [6] Department of Innovative Technologies in Medicine & Dentistry, University G. D'Annunzio, Chieti-Pescara, Italy; [7] Anatomic Pathology Unit, San Camillo-Forlanini Hospitals, Rome, Italy; [8] Medical Oncology, San Camillo Forlanini Hospital, Roma, Italy; [9] Medical Oncology, Università Cattolica del Sacro Cuore, Rome, Italy; [10] Comprehensive Cancer Center, Fondazione Policlinico Universitario Agostino Gemelli, Istituto di Ricovero e Cura a Carattere Scientifico, Rome, Italy; [11] Department of Human Pathology in Adult and Developmental Age "Gaetano Barresi", Section of Pathology, University of Messina, Messina, Italy; [12] Medical Oncology Unit, Azienda Ospedaliero Universitaria Sant'Andrea, Rome, Italy; [13] Phase 1 Unit, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy; [14] Department of Experimental Medicine, Sapienza University of Rome, Rome, Italy; [15] Pathology Unit, St. Salvatore Hospital, L'Aquila, Italy; [16] Departement de Medicine Oncologique, Institut Gustave Roussy, Villejuif, France; [17] Department of Medical, Oral, and Biotechnological Sciences University "G. D'Annunzio" of Chieti-Pescara

Received: March 22, 2024
Accepted: July 12, 2024

Correspondence
Antonio Marchetti
E-mail: amarchet@unich.it

How to cite this article: Cortellini A, Zampacorta C, De Tursi M, et al. A preliminary study on the diagnostic performance of the uPath PD-L1 (SP263) artificial intelligence (AI) algorithm in patients with NSCLC treated with PD-1/PD-L1 checkpoint blockade. Pathologica 2024;116:222-231. https://doi.org/10.32074/1591-951X-998

Summary

Objective. The uPath PD-L1 (SP263) is an AI-based platform designed to aid pathologists in identifying and quantifying PD-L1 positive tumor cells in non-small cell lung cancer (NSCLC) samples stained with the SP263 assay.

Methods. In this preliminary study, we explored the diagnostic performance of the uPath PD-L1 algorithm in defining PD-L1 tumor proportion score (TPS) and predict clinical outcomes in a series of patients with advanced stage NSCLC treated with single agent PD-1/ PD-L1 checkpoint blockade previously assessed with the SP263 assay in clinical practice.

Results. 44 patients treated from August 2015 to January 2019 were included, with baseline PD-L1 TPS of ≥ 50%, 1-49% and < 1% in 38.6%, 25.0% and 36.4%, respectively. The median uPath PD-L1 score was 6 with a significant correlation with the baseline PD-L1 TPS (r: 0.83, p < 0.01). However, only 27 cases (61.4%) were scored within the same clinically relevant range of expression (≥ vs < 50%). In the study population the baseline PD-L1 TPS was not significantly associated with clinical outcomes, while the uPath PD-L1 score showed a good diagnostic ability for the risk of death at the ROC curve analysis [AUC: 0.81 (95%CI: 0.66-0.91), optimal cut-off of ≥ 3.2], resulting in 19 patients (43.2%) being u-Path low and 25 patients (56.8%) being uPath high. The objective response rate in uPath high and low was 51.6% and 25.0% (p = 0.1), respectively, although the uPath was significantly associated with overall survival (OS, HR 2.45, 95%CI: 1.19-5.05) and progression free survival (PFS, HR 3.04, 95%CI: 1.51-6.14). At the inverse probability of treatment weighting analysis used to balance baseline covariates, the uPath categories confirmed to be independently associated with OS and PFS.

Conclusions. This preliminary analysis suggests that AI-based, digital pathology tools such as uPath PD-L1 (SP263) can be used to optimize already available biomarkers for immune-oncology treatment in patients with NSCLC.

Key words: PD-L1, digital pathology, NSCLC, Immune Checkpoint, immunotherapy.

Introduction

With the advent of programmed death/programmed death-ligand 1 (PD-1/PD-L1) immune checkpoint inhibitors (ICIs) the treatment scenario of non-small cell lung cancer (NSCLC) has radically changed [1], leading to an overall improvement of the prognosis of many patients with non-oncogene addicted disease [2]. However, only a subset of patients derives long-term benefit from these therapies, also due to the lack of high-performance predictive biomarkers.

The immunohistochemical expression of PD-L1 on tumor cells through the Tumor Proportion Score (TPS) is the only validated and routinary used biomarker [3,4], with several assays available in clinical practice, such as the 22C3 PharmDx and SP263. Even though harmonization studies demonstrated high correlation between different methodologies [5] and a certain level of interobserver reproducibility [6], several limitations have been highlighted, mainly related to inherent analytical aspects and predictive performance [7].

Current improvements in technology have produced whole slide scanners that can digitize a glass microscope slide and produce high resolution whole slide images (WSI) that can be stored and reviewed by pathologists. In addition, with digital pathology techniques, computational deep learning-based analyses can be applied to WSI to extract information and obtain quantitative data that can be used to predict outcomes [8]. Therefore, even complex patterns can be assessed by artificial intelligence (AI) based models, overcoming the limitations associated with manual scoring and human bias.

The uPath PD-L1 image analysis for NSCLC (uPath PD-L1 (SP263) AI algorithm) is designed to aid pathologists in identifying and quantifying PD-L1 positive tumor cells in NSCLC samples stained with the SP263 assay [9].

In this preliminary study, we explored the diagnostic performance of the uPath PD-L1 algorithm in defining PD-L1 TPS and predict clinical outcomes in a series of patients with advanced stage NSCLC treated with single agent PD-1/PD-L1 checkpoint blockade previously assessed with the SP263 assay in clinical practice.

Materials and methods

Study design

This is a retrospective study aimed at exploring the diagnostic ability for clinical outcomes of a new digital-pathology based technique, termed uPath PD-L1, to assess PD-L1 expression on tumor cells from formalin-fixed paraffin-embedded (FFPE) tissue, in patients with stage IV non-small cell lung cancer (NSCLC) treated with PD-1/PD-L1 checkpoint inhibitor monotherapy in clinical practice.

Patients whose PD-L1 tumor proportion score (TPS) had been assessed in clinical practice with the SP263 immunohistochemical antibody at the participating institutions (Tab. I) and treated with single agent PD-1/PD-L1 inhibitors from August 2015 to January 2019 were included in the analysis.

Table I. Patient distribution across participating centers.

Center	N (%)
University Hospital of L'Aquila, L'Aquila, Italy	19 (43.2)
University Hospital of Chieti, Chieti, Italy	10 (22.7)
Sant' Andrea University Hospital, Rome, Italy	2 (4.5)
St. Camillo-Forlanini Hospital, Rome, Italy	13 (29.5)
Total	44 (100)

The primary clinical endpoint of interest was overall survival (OS), defined as the time interval from treatment initiation to death or lost to follow-up. As secondary clinical endpoints we assessed the progression free survival (PFS), defined as the time interval from treatment initiation to disease progression or death, whichever occurred first, and objective response rate (ORR), defined as the proportion of patients experiencing an objective response (complete or partial response) as best response to treatment. Patients were assessed with radiological imaging in clinical practice, with a frequency ranging from 12 to 16 weeks; investigators were asked to provide disease assessment following the RECIST criteria v. 1.1. Patients alive at the data cut-off date were censored at the date of the last clinical follow-up for OS, while patients who did not experience disease progression by the data cut-off date were censored at the date of the last radiological assessment for PFS. The data cut-off period was September 2019.

All cases evaluated with the uPath PD-L1 system were previously centralized in a single institution and immunostained, immediately before the uPath analysis.

The TPS analysis with the uPath PD-L1 system was collectively conducted in dedicated sections by all the trained pathologists involved in the study.

After having assessed the PD-L1 tumor expression with the uPath, we explored its diagnostic performance for the risk of death with a Receiver Operating Curve (ROC) analysis, in order to establish an optimal cut-off. We then evaluated clinical outcomes using the established cut-off for uPath PD-L1 with univariable

analysis, and to provide preliminary evidence of the potentially improved diagnostic performance of the uPath based assessment of PD-L1, we explored OS and PFS according to the conventional PD-L1 TPS previously assessed in clinical practice. Finally, we validated with multivariable analysis using Inverse Probability of Treatment Weighing (ITPW) procedures the differential effect of conventional PD-L1 TPS assessment and the uPath PD-L1 assessment in determining clinical outcomes.

PD-L1 expression

Being an observational retrospective study with no intervention planned, only patients whose PD-L1 TPS had been previously assessed in clinical practice with the SP263 immunohistochemical antibody were considered eligible. In addition, to ensure reproducibility of our preliminary findings, we restricted the analysis to patients with previously assessed PD-L1 TPS from FFPE tissue (either from the primary tumor or metastatic sites), while patients with only cytological samples available were excluded.

Five µm sections immunostained with Ventana SP263 antibody on the Ultra benchmark platform were digitized with the Ventana DP 200 scanner, ensuring high-resolution images at 40x magnification. These digital slides were then uploaded onto the uPath platform for analysis, leveraging a specialized algorithm designed to calculate the PD-L1 expression percentage [9]. This AI algorithm is trained by a pathologist and provides a reliable and reproducible target score for scanned images[9]. Considering that the software offers the possibility to analyze the entire slide before pathologist's evaluation, or to provide results for each pathologist-selected area of interest, referred to as Region Of Interest (ROI) [9], we opted to specifically apply the algorithm to ROI on the basis of a preliminary analysis suggesting more accurate results (data not shown). For example, on the basis of each sample features, we excluded necrotic or suppurative areas that might have potentially affected TPS values. In addition, it computes the average score across all ROIs. Once the analysis is complete, the software visually distinguishes between positively and negatively stained tumor cells by marking them with red and blue dots, respectively [9].

Statistical analysis

Descriptive statistics were used to report baseline clinic-pathologic characteristics. The chi-square test was used to determine associations between categorical variables. Pearson's correlation coefficient (r) was used to test the correlation between PD-L1 TPS previously assessed in clinical practice and the uPath PD-L1 TPS value. The median period of follow-up was computed with the reverse Kaplan-Meier method. A Cox proportional hazard regression was used to compute the hazard ratios (HR) with 95% confidence intervals (CIs) and the predicted probabilities for death using the uPath PD-L1 as a continuous independent variable. Subsequently, we used a ROC curve analysis to compute the area under the curve (AUC) for the risk of death according to the uPath PD-L1 predicted probabilities and determined an optimal cut-off for survival using Youden's J statistic.

Univariable analyses of OS and PFS were performed using the Kaplan-Meier method and the log-rank test. Considering the limited sample size, IPTW procedures to mitigate the potentially different distribution of patient's characteristics according to the conventional PD-L1 TPS subgroups (≥50% vs < 50%) and the uPath PD-L1 subgroups (high vs low), in order to run IPTW-fitted analyses for each clinical outcome. The included covariates were: age (≥ vs < 70 years), gender (male vs female), histology (squamous vs non-squamous), smoking status (ever smoker vs smoker), Eastern Cooperative Oncology Group Performance Status (ECOG-PS) (0-1 vs ≥ 2), number of metastatic sites (> 2 vs ≤ 2) and treatment line (first vs non-first). Lastly, considering that the data source consisted of four different centers, and that baseline PD-L1 TPS assessment can be operator-dependent, a conditional interpretation for participating center by using frailty models was applied to correct all the 95%CI of Cox regressions.

Analyses were performed using the R-studio software, R Core Team 2021 (R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria), and the MedCalc® Statistical Software version 20 (MedCalc Software Ltd, Ostend, Belgium; https://www.medcalc.org; 2021).

Results

Cohort characteristics

Overall, 44 patients were included in the analysis, with a median age of 67 years (range: 39-89). As reported in Table II, the majority of patients were males (31, 70.5%), ever smokers (31 (70.5%), had non-squamous histology (23, 52.3%), presented a baseline ECOG-PS of 0-1 (32, 72.7%), had ≤ 2 metastatic sites and were treated in the advanced line setting (29, 65.9%). Baseline PD-L1 TPS were ≥ 50% in 17 patients (38.6%), 1-49% in 11 patients (25.0%) and < 1% in 16 patients (36.4%), with a median value of 15% (range: < 1%-100%, IQR: < 1%-60%). The

Table II. Baseline patient characteristics.

Variable	Overall 44 N (%)
Age	
Median (range)	67 (39-89)
< 70 years	26 (59.1)
≥ 70 years	18 (40.9)
Gender	
Female	13 (29.5)
Male	31 (70.5)
Histology	
Non-Squamous	23 (52.3)
Squamous	21 (47.7)
Smoking status	
Ever smokers	31 (70.5)
Never smokers	13 (29.5)
ECOG-PS	
0-1	32 (72.7)
≥ 2	12 (27.3)
Number of metastatic sites	
≤ 2	25 (56.8)
> 2	19 (43.2)
Treatment line	
First	15 (34.1)
Non-first	29 (65.9)
Treatment regimen	
Nivolumab	27 (61.4)
Pembrolizumab	17 (38.6)
PD-L1 TPS	
≥ 50%	17 (38.6)
1-49%	11 (25.0)
< 1%	16 (36.4)

ECOG-PS, Eastern Cooperative Oncology Group Performance Status; PD-L1: programmed death ligand 1; TPS: tumor proportion score.

baseline TPS assigned with the SP263 test was blindly recalculated by the panel of pathologist involved in the study. No major discrepancies affecting treatment decision were observed.

PD-L1 EXPRESSION ANALYSIS

After PD-L1 image digitalization the uPath algorithm was applied to all the 44 patients of the cohort resulting in a median uPath PD-L1 score of 6 (range: 0-88.2; IQR: 0.9-38.5) and a significant correlation with the baseline PD-L1 TPS (r: 0.83, 95%CI: 0.71-0.91, p < 0.0001, Tab. III). Applying the standard cut-off of 50%, the results of the analysis are consistent with those obtained with the conventional method in 61.4% of cases, for a total of 27 cases falling within the same range of expression (≥ 50% vs < 50%), being therefore potentially eligible for the same front line, ICI-based treatment with both analytical approaches. However, in 17 (38.6%) of cases, a discrepancy was found between the two techniques. The two most frequent situations in discordant cases were the following: cases with low expression of PD-L1 with a change from a value < 1% with conventional method to one within the range

Table III. Paired samples for previously assessed PD-L1 TPS and the uPath PD-L1 score. PD-L1: programmed death ligand 1: TPS: tumor proportion score.

Sample ID	PD-L1 TPS	u-Path PD-L1
1	100	88.2
2	100	50.8
3	100	49.2
4	100	40
5	95	81
6	95	56.4
7	95	47.5
8	80	84.5
9	80	66.9
10	80	21.7
11	60	80.9
12	60	22.5
13	60	9.9
14	52	23.6
15	50	69
16	50	37.1
17	50	21.7
18	40	25.4
19	40	8.6
20	40	2.7
21	35	0.6
22	15	8.6
23	15	1.5
24	10	5.4
25	5	3.2
26	5	2
27	5	1.7
28	5	1.4
29	0	7.1
30	0	6.6
31	0	3.7
32	0	2.4
33	0	1
34	0	1
35	0	0.8
36	0	0.7
37	0	0.5
38	0	0.5
39	0	0.5
40	0	0.4
41	0	0.1
42	0	0.1
43	0	0
44	0	0

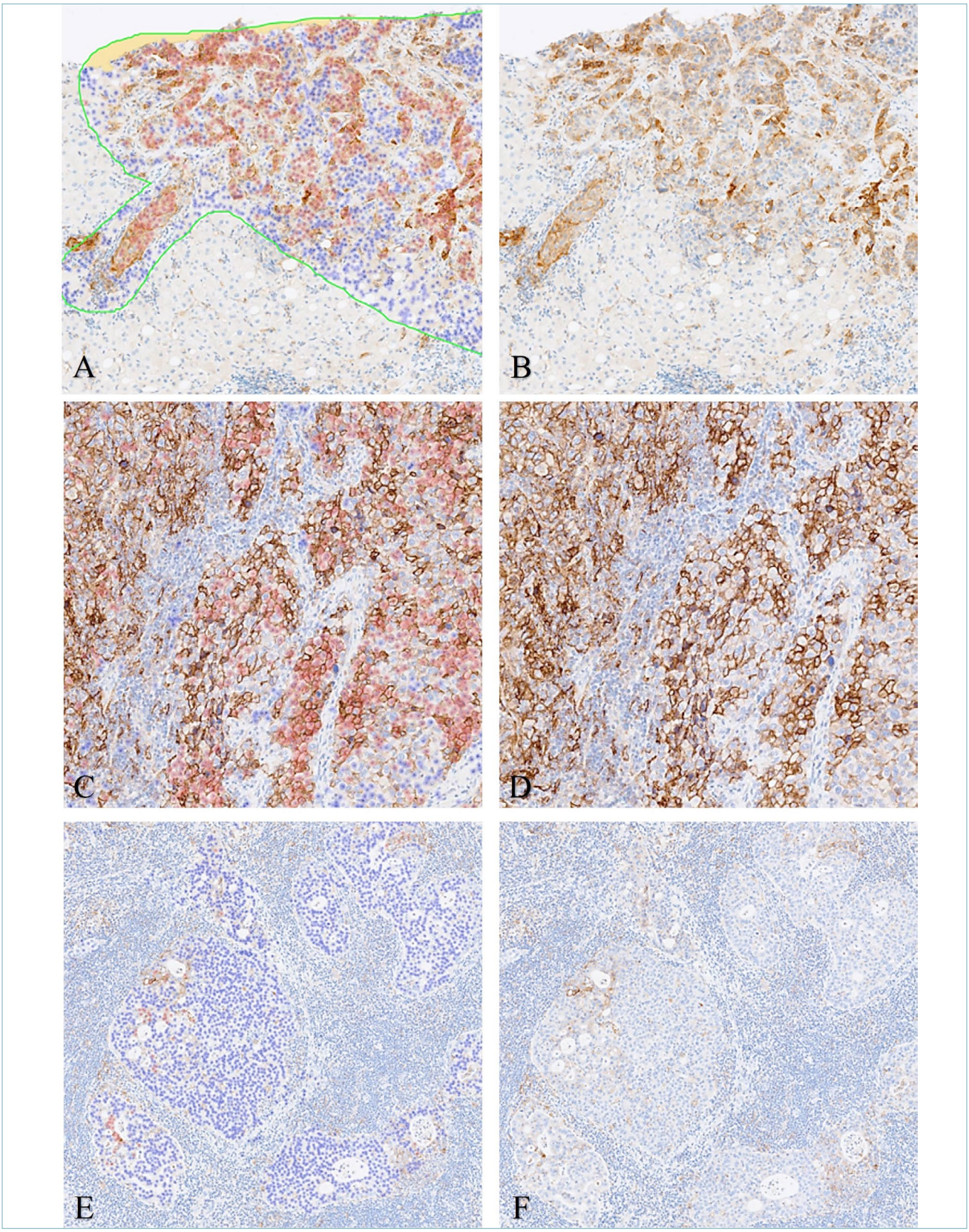

Figure 1. Examples of paired PD-L1 (anti PD-L1 SP263) images analyzed with the uPath algorithm and conventional tumor proportion score (TPS). Liver metastasis from lung squamous cell carcinoma, uPath assessment: 56.4%, with green line of demarcation of the ROI (A) and conventional assessment: 95% (B). Poorly differentiate lung cell carcinoma, uPath assessment: 66.9% (C) and conventional assessment: 80% (D). Squamous cell carcinoma, uPath assessment: 1.4% (E) and conventional assessment: 5% (F).

1-49% with the uPath algorithm; cases with intermediate/high expression with transition from a value > 50% with conventional analysis to one within the range 1-49% with the uPath algorithm (Tab. III). Some examples of PD-L1 images analyzed with the conventional method and the uPath algorithm are provided in Figure 1.

Clinical outcomes

The median follow-up for the entire cohort was 18.2 months (95%CI: 15.0-28.2). When used as a continuous variable, the uPath PD-L1 score was not significantly associated with the risk of death (HR 0.99, 95%CI: 0.97-1.01). We subsequently used the predicted probability of death for each uPath PD-L1 value to run a ROC curve analysis for the risk of death, which showed a good diagnostic ability of the uPath PD-L1 score with an AUC of 0.81 (95%CI: 0.66-0.91; p < 0.0001, Fig. 2A) and an optimal cut-off set at ≥ 3.2, resulting in 19 patients (43.2%) being u-Path low and 25 patients (56.8%) being uPath high.

When used as a categorical variable, the uPath score was not associated with tumor response as the ORR among the uPath low and uPath high group were 25.0% (95%CI: 3.0-90.3) and 51.6% (95%CI: 29.5-83.8) (p = 0.1839, Fig. 2B). However, uPath was significantly associated with OS (HR 2.45, 95%CI: 1.19-5.05, Fig. 2C) and PFS (HR 3.04, 95%CI: 1.51-6.14, Fig. 2D).

We then assessed OS and PFS according to the baseline PD-L1 TPS with univariable analysis in the study cohort as summarized in Figure 3 A-D, with no

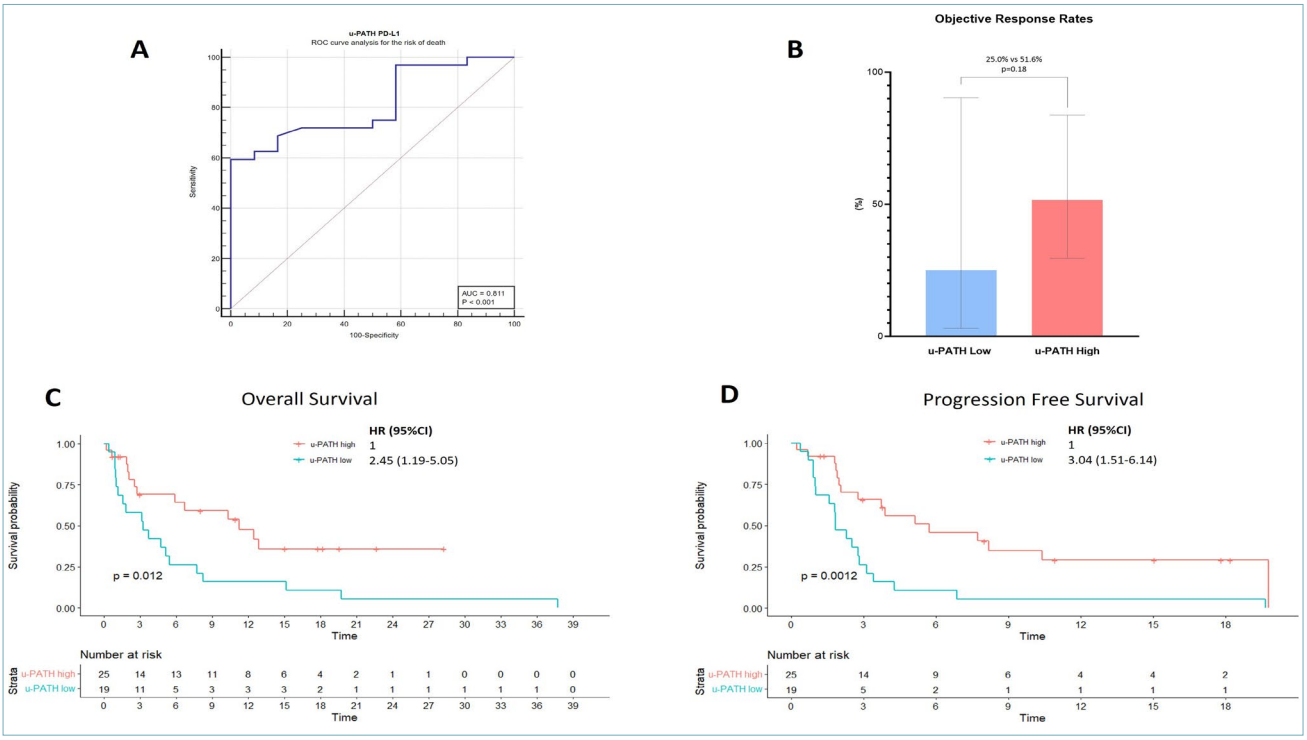

Figure 2. uPath PD-L1 diagnostic performance summary. A) ROC curve analysis for the risk of death: with 32 events, the AUC was 0.81 (95%CI: 0.66-0.91; p < 0.0001), with a sensitivity of 59.8 and a specificity of 100.0. The Youden index J was 0.59 and was associated at the uPath PD-L1 score of > 3.2. B) ORR analysis: 2 disease responses were observed among 8 evaluable patients with uPath low tumors, while 16 disease responses were observed among 31 evaluable patients with uPath high tumors. C) OS Kaplan-Meier estimates: patients with uPath low tumors achieved an OS of 3.2 months (95%CI: 1.57-8.24, 19 events) while patients with uPath high tumors achieved an OS of 11.2 months (95%CI: 5.9-not reached, 13 events). Log-rank p-value: 0.012. HR: 2.45 (95%CI: 1.19-5.05). D) PFS Kaplan-Meier estimates: patients with uPath low tumors achieved a PFS of 1.8 months (95%CI: 1.57-3.4, 16 events) while patients with uPath high tumors achieved a PFS of 5.7 months (95%CI: 2.8-not reached, 19 events). Log-rank p-value: 0.001. HR: 3.45 (95%CI: 1.51-6.14). A conditional interpretation for participating center using frailty models was applied to correct all the 95%CI from Cox regressions. ROC: receiver operating characteristic; AUC: area under the curve; ORR: objective response rate; OS: overall survival; PFS: progression free survival; HR: hazard ratio; 95%CI: 95% confidence interval.

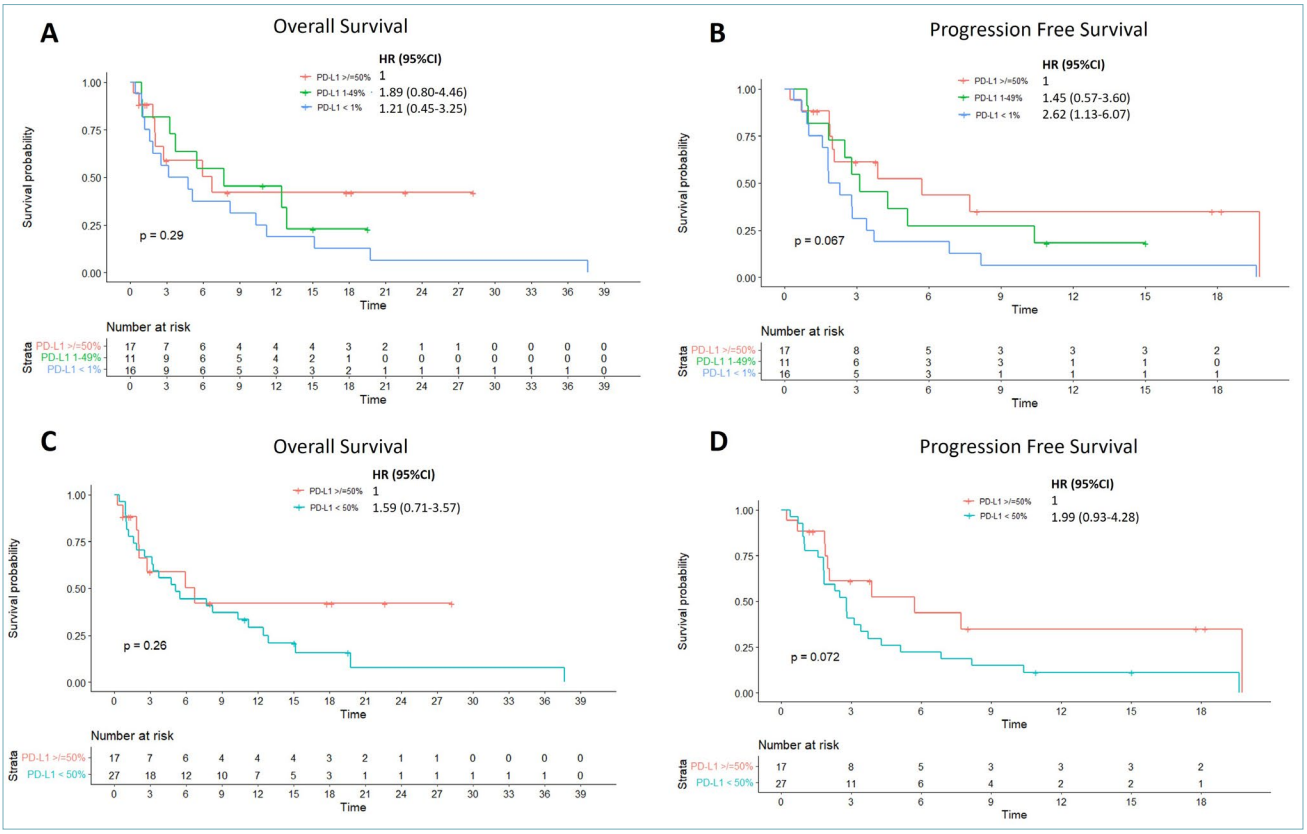

Figure 3. Kaplan-Meier survival estimates according to previously assessed PD-L1 TPS. A) OS: patients with PD-L1 < 1% an OS of 3.9 months (95%CI: 1.5-15.2, 16 events), patients with PD-L1 1-49% achieved an OS of 7.7 months (95%CI: 3.7-not reached, 8 events), and patients with PD-L1 ≥50% achieved an OS of 6.7 months (95%CI: 2.1-not reached, 8 events). Log-rank p-value: 0.29. B) PFS: patients with PD-L1 < 1% achieved a PFS of 2.1 months (95%CI: 1.5-6.9, 16 events), patients with PD-L1 1-49% achieved a PFS of 3.1 months (95%CI: 2.5-not reached, 9 events), and patients with PD-L1 ≥50% achieved a PFS of 5.7 months (95%CI: 1.9-not reached, 10 events). Log-rank p-value: 0.07. C) OS: patients with PD-L1 < 50% achieved an OS of 5.1 months (95%CI: 3.1-12.5, 24 events), patients with PD-L1 ≥50% achieved an OS of 6.7 months (95%CI: 2.1-not reached, 8 events). Log-rank p-value: 0.26. D) PFS: patients with PD-L1 < 50% achieved a PFS of 2.8 months (95%CI: 1.8-4.3, 25 events), patients with PD-L1 ≥50% achieved a PFS of 5.7 months (95%CI: 1.9-not reached, 10 events). Log-rank p-value: 0.072. A conditional interpretation for participating centre using frailty models was applied to correct all the 95%CI from Cox regressions. PD-L1 programmed death ligand-1; TPS: tumor proportion score; OS: overall survival; PFS: progression free survival; HR: hazard ratio; 95%CI: 95% confidence interval.

Table IV. Distribution of baseline characteristics before and after the IPTW procedure between patients with u-Path high and u-Path low tumors. Variability of included characteristics is estimated through the standardized mean difference (SMD).

	u-PATH high (%)	u-PATH low (%)	P value	u-PATH high Weighted (%)	u-PATH low Weighted (%)	P value	SMD
Smoking status							
Never smokers	44	10.5	0.03	36.5	17.8	0.33	0.43
Sex							
Male	64.0	78.9	0.45	62.8	61.7	0.95	0.02
Age							
≥ 70 years old	48.0	31.6	0.43	51.2	56.6	0.77	0.10
ECOG-PS							
≥ 2	12.0	47.4	0.02	15.5	45.7	0.09	0.69
Number of metastatic sites							
> 2	40.0	47.4	0.85	37.1	32.8	0.79	0.09
Treatment line							
Non-first	48.0	89.5	0.01	55.0	73.3	0.43	0.38
Histology							
Non-squamous	48.0	47.4	1.0	45.8	42.9	0.87	0.05

ECOG-PS: easter cooperative oncology group-performance status; IPTW: inverse probability of treatment weighing.

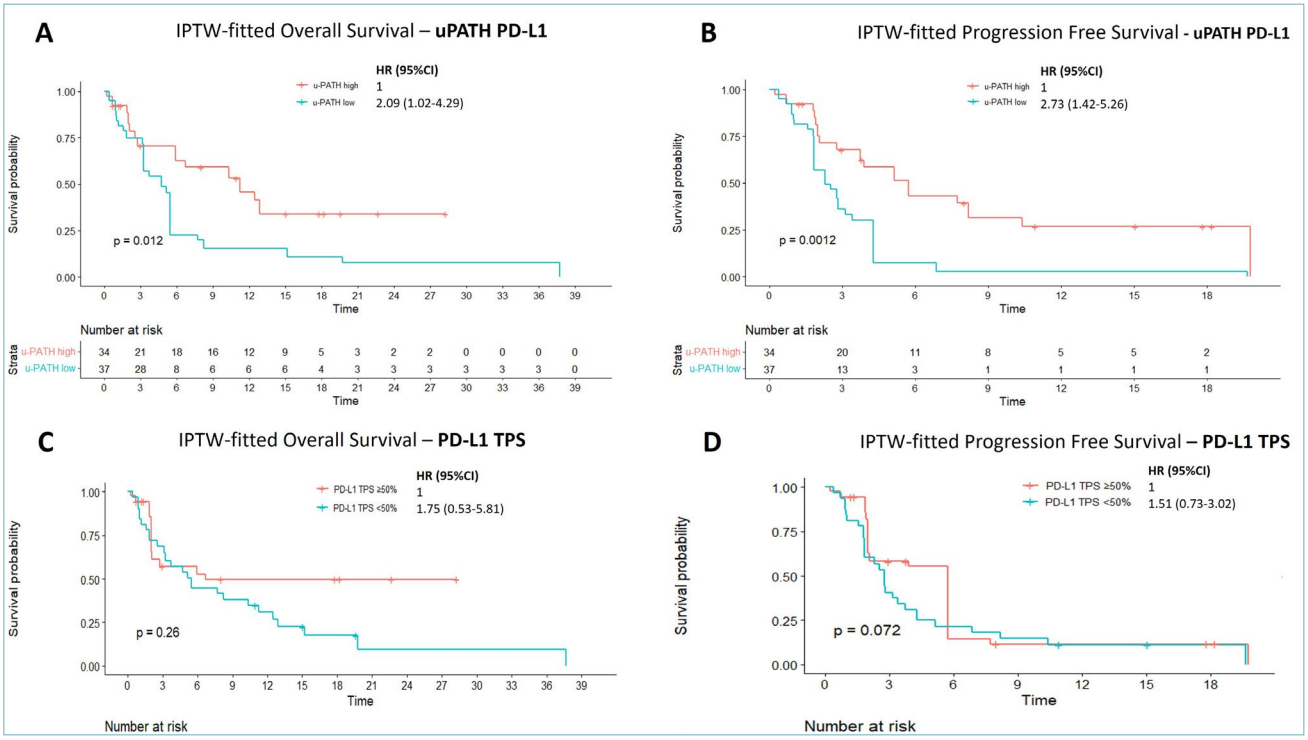

Figure 4. IPTW-fitted Kaplan-Meier survival estimates according to uPath PD-L1 assessment. A) OS analysis: patients with uPath low tumors achieved an OS of 4.7 months (95%CI: 3.2-15.2) while patients with uPath high tumors achieved an OS of 11.2 months (95%CI: 5.9-not reached). Log-rank p-value: 0.012. HR: 2.09 (95%CI: 1.02-4.29). B) PFS analysis: patients with uPath low tumors achieved a PFS of 2.3 months (95%CI: 1.8-4.3), while patients with uPath high tumors achieved a PFS of 5.7 months (95%CI: 3.7-not reached). Log-rank p-value: 0.001. HR: 2.73 (95%CI: 1.42-2.56). C) OS analysis: patients with PD-L1 TPS < 50% achieved an OS of 5.4 months (95%CI: 3.1-12.9), while patients with PD-L1 TPS ≥50% achieved an OS of 6.7 months (95%CI: 1.9-not reached). Log-rank p-value: 0.26. HR: 1.75 (95%CI: 0.53-5.81). D) PFS analysis: patients with PD-L1 TPS < 50% achieved a PFS of 2.8 months (95%CI: 1.8-4.3), while patients with PD-L1 TPS ≥50% achieved a PFS of 5.7 months (95%CI: 1.9-5.7). Log-rank p-value: 0.072. HR: 1.51 (95%CI: 0.75-3.02). A conditional interpretation for participating center using frailty models was applied to correct all the 95%CI from Cox regressions. PD-L1 programmed death ligand-1; OS: overall survival; PFS: progression free survival; HR: hazard ratio; 95%CI: 95% confidence interval.

Table V. Distribution of baseline characteristics before and after the IPTW procedure between patients with PD-L1 TPS ≥50% and PD-L1 TPS < 50% tumors. Variability of included characteristics is estimated through the standardized mean difference (SMD). ECOG-PS: easter cooperative oncology group-performance status; IPTW: inverse probability of treatment weighing.

	PD-L1 ≥50% (%)	u PD-L1 < 50% (%)	P value	PD-L1 ≥50% Weighted (%)	u PD-L1 < 50% Weighted (%)	P value	SMD
Smoking status							
Never smokers	52.9	14.8	0.018	30.1	18.9	0.49	0.26
Sex							
Male	76.5	66.7	0.723	52.8	64.8	0.61	0.24
Age							
≥ 70 years old	41.2	40.7	1.00	42.1	45.0	0.89	0.05
ECOG-PS							
≥ 2	17.6	33.3	0.43	27.3	37.6	0.55	0.22
Number of metastatic sites							
> 2	52.9	37.0	0.46	10.9	33.1	0.10	0.55
Treatment line							
Non-first	23.5	92.6	< 0.01	65.0	88.1	0.10	0.56
Histology							
Non-squamous	41.2	51.9	0.70	35.9	49.3	0.54	0.27

significant association when used as a three-category variable (≥ 50% vs 1-49% vs < 1%) nor when used as a dichotomous variable (≥ 50% vs < 50%).
To minimize the differential distribution of baseline characteristics according to the uPath PD-L1 stratification, we performed an IPTW procedure as reported in Table IV**.** The IPTW-fitted analysis eventually confirmed the u-Path PD-L1 as an independent predictor of OS (HR 2.09, 95%CI: 1.02-4.29, Fig. 4A) and PFS (HR 2.73, 95%CI: 1.42-5.26, Fig. 4B). Table V reports the IPTW stratification according to the conventional PD-L1 TPS category, with the IPTW-fitted analysis according to the conventional PD-L1 showing no significant impact on clinical outcomes (Figs. 4C and 4D).

Discussion

Response and clinical benefit to ICI therapy is defined by a variety of intertwined factors including tumoral and host determinants. The treatment landscape of NSCLC is constantly evolving and additional predictive biomarkers beyond PD-L1 expression have been the focus of intense research [10], while AI-based techniques are increasingly used to optimize already available information and inform clinical decisions, including digital/AI-based pathology technologies [11].
In this study, we explored the diagnostic performance for clinical outcomes of the uPath PD-L1 algorithm in a cohort of advanced stage NSCLC treated with single agent PD-1/PD-L1 inhibitors in comparison to the conventional TPS method. Despite the significant correlation (r: 0.83) between the two methodologies, we found inconsistencies using pre-established PD-L1 categories, partially colliding with previously reported data using other AI-assisted PD-L1 estimation techniques [12]. Using the 50% cut-off, only 61.4% of the cases fell within the same diagnostic range (≥ vs < 50%), while 38.6% of cases showed mixed results, with important potential implications for the first-line single agent vs combination chemo-immunotherapy eligibility.
These discrepancies resemble a recent study, where although the implementation of the uPath algorithm improved the concordance of assessing PD-L1 expression between pathologists, a shift of the clinically relevant category (negative/weak/strong expression) was reported for several cases [13]. Inter-test inconsistencies can occur because the algorithm has a better computational capacity than a pathologist, allowing quick and accurate enumeration of tumor cells on the digitalized slide, especially in low ranges of expression (between 1 and 10%). This could explain scenarios in which the TPS data has changed from a value < 1% with the conventional method to one within the range 1-49% with the AI algorithm.
In addition, the AI algorithm has shown remarkable precision in detecting and assessing membrane positivity. Traditional pathologist analysis can be labor-intensive, especially in instances with heterogeneous PD-L1 expression and in cases involving slides obtained from surgical samples. These instances require the evaluation of large sections of tumor tissue at high magnification. Under a light microscope, it can often be challenging to differentiate cytoplasmic staining, which should not be included in the TPS calculation, from membrane immunoreactivity which should be counted. For cases showing high expression ranges (≥ 50%) this may lead a pathologist to overestimate the number of positive tumor cells and assign a higher TPS value than what is calculated with the AI algorithm. In relation to this, the data obtained through AI tends to display TPS values that are lower than those conventionally evaluated, particularly for cases with intermediate or high expression. This discrepancy could have resulted in variations from a TPS value greater than 50% with conventional analysis to a value within the 1-49% range using the AI algorithm.
The greater accuracy of the uPath PD-L1 is reflected in the survival analysis, as the conventional PD-L1 TPS was not significantly associated with clinical outcomes in the included cohort, whilst the uPath PD-L1 score showed a significant diagnostic ability for the risk of death (AUC: 0.81). In addition, with the preliminary established cut-off of ≥ 3.2% the uPath PD-L1 showed to be significantly associated with OS and PFS even in IPTW-fitted models, while uPath PD-L1 high showed numerically increased tumor response in comparison to uPath PD-L1 low patients (51.6% vs 25.0%).
In keeping with our results, a recent study have confirmed a better TPS evaluation through the use of AI assistance compared to conventional analysis alone and, therefore, a better prediction of the therapeutic response [14].
Our study acknowledges several limitations, mainly deriving from its retrospective design and the small sample size, which did not allow us to make a fully powered analysis for clinical outcomes and to properly assess the uPath PD-L1 score through separate training and validation cohorts. The limited sample size additionally limited our ability to explore the uPath PD-L1 score as a continuous variable. However, this exploratory analysis sums up to the accumulating evidence indicating the possible implementation of AI-based techniques in modern day pathology, to optimize currently available biomarkers.

Conclusions

Our preliminary study indicates the uPath PD-L1 (SP263) AI algorithm is a potential tool to improve the diagnostic ability of the PD-L1 TPS in patients with advanced stage NSCLC treated with PD-1/PD-L1 checkpoint blockade and supports further research efforts to fully validate its use in clinical practice.

Acknowledgements

Roche Diagnostics kindly supplied the Ventana DP200 Scanner with relative software and the algorithm license.

Conflicts of interest statement

The authors disclose no conflicts of interest in relation to the published work.
Alessio Cortellini received grants for consultancies/advisory boards: BMS, MSD, OncoC4, IQVIA, Roche, GSK, AstraZeneca, Access Infinity, Ardelis Health and REGENERON. He also received speaker fees from AstraZeneca, EISAI, MSD, SANOFI/REGENERON and Pierre-Fabre. All other authors declare no conflict of interest.

Funding

This study was partially funded by the National Operational Program (PON) ARS01_01195 from the Italian Ministry of Education, University and Research (MIUR).

Authors' contributions

All authors contributed to the publication according to the ICMJE guidelines for the authorship (study conception and design, acquisition of data, analysis and interpretation of data, drafting of manuscript, critical revision). All authors read and approved the submitted version of the manuscript (and any substantially modified version that involves the author's contribution to the study). Each author has agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Ethical consideration

The procedures followed were in accordance with the precepts of Good Clinical Practice and the declaration of Helsinki. The study was approved by the respective local ethical committees on human experimentation of each institution, after previous approval by the coordinating center (Comitato Etico per le province di Chieti e Pescara, IRB ID approval N.25, 6 December 2018).

References

1 Desai A, Peters S. Immunotherapy-based combinations in metastatic NSCLC. Cancer Treat Rev. 2023/05/01/ 2023;116:102545. https://doi.org/10.1016/j.ctrv.2023.102545

2 Berghmans T, Dingemans A-M, Hendriks LEL, et al. Immunotherapy for nonsmall cell lung cancer: a new therapeutic algorithm. Eur Respir J. 2020;55(2):1901907. https://doi.org/10.1183/13993003.01907-2019

3 Garon EB, Rizvi NA, Hui R, et al. Pembrolizumab for the Treatment of Non–Small-Cell Lung Cancer. N Engl J Med. 2015;372(21):2018-2028. https://doi.org/10.1056/NEJMoa1501824

4 Reck M, Rodríguez-Abreu D, Robinson AG, et al. Pembrolizumab versus Chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer. N Engl J Med. 2016;375(19):1823-1833. https://doi.org/10.1056/NEJMoa1606774

5 Marchetti A, Barberis M, Franco R, et al. Multicenter Comparison of 22C3 PharmDx (Agilent) and SP263 (Ventana) Assays to Test PD-L1 Expression for NSCLC Patients to Be Treated with Immune Checkpoint Inhibitors. J Thorac Oncol. 2017;12(11):1654-1663. https://doi.org/10.1016/j.jtho.2017.07.031

6 Chang S, Park HK, Choi Y-L, Jang SJ. Interobserver Reproducibility of PD-L1 Biomarker in Non-small Cell Lung Cancer: A Multi-Institutional Study by 27 Pathologists. J Pathol Transl Med. 11 2019;53(6):347-353. https://doi.org/10.4132/jptm.2019.09.29

7 Teixidó C, Vilariño N, Reyes R, et al. PD-L1 expression testing in non-small cell lung cancer. Ther Adv Med Oncol. 2018;10:1758835918763493. https://doi.org/10.1177/1758835918763493

8 Gupta R, Kurc T, Sharma A, et al. The Emergence of Pathomics. Current Pathobiology Reports. 2019/09/01 2019;7(3):73-84. https://doi.org/10.1007/s40139-019-00200-x

9 uPath PD-L1 (SP263) image analysis, NSCLC*. Accessed 19 Dec 2023, https://diagnostics.roche.com/global/en/products/digital/upath-pd-l1-sp263-image-analysis-nsclc-pid00000381.html

10 Bodor JN, Boumber Y, Borghaei H. Biomarkers for immune checkpoint inhibition in non–small cell lung cancer (NSCLC). Cancer. 2020;126(2):260-270. https://doi.org/10.1002/cncr.32468

11 Prelaj A, Miskovic V, Zanitti M, et al. Artificial intelligence for predictive biomarker discovery in immuno-oncology: a systematic review. Ann Oncol. 2024;35(1):29-65 https://doi.org/10.1016/j.annonc.2023.10.125

12 Wu J, Liu C, Liu X, et al. Artificial intelligence-assisted system for precision diagnosis of PD-L1 expression in non-small cell lung cancer. Mod Pathol. 2022;35(3):403-411. https://doi.org/10.1038/s41379-021-00904-9

13 Haragan A, Parashar P, Bury D, et al. Machine-learning-based image analysis algorithms improve interpathologist concordance when scoring PD-L1 expression in non-small-cell lung cancer. J Clin Pathol. 2023:jcp-2023-208978. https://doi.org/10.1136/jcp-2023-208978

14 Choi S, Cho SI, Ma M, et al. Artificial intelligence-powered programmed death ligand 1 analyser reduces interobserver variation in tumour proportion score for non-small cell lung cancer with better prediction of immunotherapy response. Eur J Cancer. 2022;170:17-26. https://doi.org/10.1016/j.ejca.2022.04.011. Epub 2022 May 14. PMID: 35576849.