Assessment of Concordance between 22C3 and SP142 Immunohistochemistry Assays regarding PD-L1 Expression in Non-Small Cell Lung Cancer

Haipeng Xu1, Gen Lin2, Cheng Huang1, Weifeng Zhu2, Qian Miao1, Xirong Fan1, Biao Wu1, Xiaobing Zheng1, Xiandong Lin3, Kan Jiang1, Dan Hu2 & Chao Li2,4,5

Different anti-PD-1 and anti-PD-L1 antibodies bind different epitopes. However, whether the results from the SP142 and 22C3 immunochemistry (IHC) assays can be interchanged to determine patient eligibility for immunotherapy remains largely unknown. Histologic sections from 135 tumor samples were probed with both 22C3 and SP142 antibodies. The concordance of PD-L1 expression determined by the two assays was assessed. Additionally, we evaluated the association of PD-L1 expression detected by different assays with clinicopathological features and prognosis. In total, 105 (77.78%) of 135 samples evaluated by the 22C3-IHC platform produced the same results with the SP142-IHC platform (Kappa value: 0.481, \( p < 0.001 \)). In addition, 69 (51.11%) of 135 samples evaluated by the SP142-IHC platform produced the same results with the 22C3-IHC platform (Kappa value: 0.324, \( p < 0.001 \)). PD-L1 expression based on the 22C3-IHC assay was significantly correlated with smoking status, whereas that based on the SP142-IHC assay was correlated with smoking status, sex, and histology. Compared to the SP142-IHC assay, the 22C3-IHC assay usually resulted in an underestimation of PD-L1 expression in tumor cells and immune cells. Thus, the results from the two assays cannot be interchanged. Our data also suggest that the use of different reagents may account for inconsistencies in the literature regarding the association between PD-L1 expression and clinicopathological features.

Programmed death ligand-1 (PD-L1) expression is a major immune-suppressive mechanism that is stimulated by the engagement of the PD-1/PD-L1 axis in non-small cell lung cancer (NSCLC). Tumor cells (TCs) can evade immune responses through the upregulation of PD-L1 and blockade of the PD-1/PD-L1 interaction via monoclonal antibodies and can produce a durable clinical response in patients with NSCLC. The anti-PD-1 antibody pembrolizumab and the anti-PD-L1 antibody atezolizumab are now approved by the US Food and Drug Administration (FDA) for treating patients with advanced NSCLC1,2.

There is data indicating that the PD-L1 expression status might help guide therapy3,4. For example, the results from the phase II/III KEYNOTE-010 study showed that compared with docetaxel, pembrolizumab prolonged overall survival (OS) in previously treated advanced NSCLC patients with positive PD-L1 expression determined by immunohistochemistry (IHC) on at least 1% of TCs5. Data from patients with advanced NSCLC in a phase III study (KEYNOTE-024) demonstrated an objective response rate (ORR) with pembrolizumab in 44.8% of patients with at least 50% PD-L1-positive TCs compared with the ORR of 27.8% with chemotherapy. The median progression-free survival (PFS) was 10.3 months in the pembrolizumab group versus 6.0 months in the

1Department of Thoracic Oncology, Fujian Cancer Hospital, Fujian Medical University Cancer Hospital, Fuzhou, 350014, China. 2Department of Pathology, Fujian Cancer Hospital, Fujian Medical University Cancer Hospital, Fuzhou, 350014, China. 3Department of Molecular Pathology, Fujian Cancer Hospital, Fujian Medical University Cancer Hospital, Fuzhou, 350014, China. 4Fujian Provincial Key Laboratory of Translational Cancer Medicine, Fuzhou, 350014, China. 5Department of Pathology, School of Basic Medical Sciences, Fujian Medical University, Fuzhou, 350014, China. Correspondence and requests for materials should be addressed to C.L. (email: lichao3501@163.com)
PD-L1 expression was scored at three levels, namely, the Proportion Score (TPS), which is the percentage of viable TCs showing partial or complete membrane staining. For the 22C3-IHC assay, PD-L1 expression was determined by using a Tumor Proportion Score (TPS) cut-off value, and the staining intensity was not part of this evaluation. For the SP142-IHC assay standards were compared with those of PD-L1-negative tumors using the \( \chi^2 \) test. The OS of patients with PD-L1-positive and PD-L1-negative tumors was estimated using the Kaplan–Meier method, and differences were compared using the log-rank test. All tests were two-sided. Statistical significance was set at \( p < 0.05 \). Statistic analyses were performed using SPSS 21.0 software.

Results

Patient Characteristics. A total of 135 patients were eligible for the study, and their characteristics are summarized in Table 1. Ninety-eight of the 135 patients were male (72.6%). The ECOG performance status was 0 in all patients. Adenocarcinoma and squamous carcinoma were the major histologic subtypes, accounting for 39.3% and 49.6%, respectively. One hundred and twenty patients (96.3%) had stage I-III disease, and 15 (3.7%) had stage IV disease.
Comparison of PD-L1 Expression Results Based on the 22C3-IHC and SP142-IHC Assays. PD-L1 staining was observed in both TCs and ICs by both 22C3-IHC and SP142-IHC assays. Representative examples of PD-L1 staining in TCs and ICs based on two assays are shown in Fig. 1. Compared with the 22C3-IHC assay, the SP142-IHC assay resulted in weaker staining of TC membranes and fewer positive TCs (Fig. 2). One sample demonstrated 1% PD-L1-positive TCs based on the SP142-IHC assay, but it was negative based on the 22C3-IHC assay. The variability in the staining pattern of ICs between samples was greater than that for TCs (Figs 2 and 3). In a total of 62 discordant samples, compared to the 22C3-IHC assay, the SP142-IHC assay resulted in the underestimation and overestimation of PD-L1 expression in ICs in 53 (39.2%) and 9 (6.7%) samples, respectively (p < 0.001). With the adoption of the 22C3-IHC assay scoring algorithm for TCs, 20 (14.8%), 25 (18.5%), and 90 (66.7%) of the 135 samples were classified into groups that were strongly positive, weakly positive, and negative.

Table 1. Clinicopathological Characteristics of Patients.

| Patient demographics (N = 135) | N (%) |
|-------------------------------|-------|
| Age                           |       |
| >60 y                         | 62 (26.3%) |
| ≤60 y                         | 83 (73.7%) |
| Sex                           |       |
| Male                          | 98 (72.6%) |
| Female                        | 37 (27.4%) |
| Smoking status                |       |
| Never smoker                  | 72 (55.8%) |
| Former/Current smoker         | 57 (44.2%) |
| Histology                     |       |
| Squamous carcinoma            | 53 (39.3%) |
| Adenocarcinoma                | 67 (49.6%) |
| Other pathological types      | 15 (11.1%) |
| TNM stage                     |       |
| I                             | 62 (45.9%) |
| II                            | 27 (20.0%) |
| III                           | 41 (30.4%) |
| IV                            | 5 (3.7%) |

Figure 1. Representative micrographs of two PD-L1 IHC assays for TCs or ICs. (A) 22C3-IHC (80%) and SP142-IHC (70%) assays indicating strong PD-L1-positivity in TCs; (B) 22C3-IHC (10%) and SP142-IHC (10%) assays indicating strong PD-L1-positivity in ICs; (C) 22C3-IHC (30%) and SP142-IHC (5%) assays indicating weak PD-L1-positivity in TCs; and (D) 22C3-IHC (10%) assay indicating strong PD-L1-positivity in ICs, whereas SP142-IHC assay indicative negative staining in ICs.
for PD-L1, respectively. Moreover, 105 (77.78%) of 135 samples produced the same results with the SP142-IHC reagent assay (κ value: 0.481, p < 0.001). The PD-L1 levels on TCs as detected by the 22C3-IHC assay were underestimated by the SP142-IHC assay in 29 samples. Only one sample with negative PD-L1 expression based on the 22C3-IHC assay was misclassified as having weak positive PD-L1 expression based on the SP142-IHC assay (McNemar-Bowker analysis: p = 0.001) (Table 2). When using 50% PD-L1-positive TCs as the cut-off value, more than 50% of cells showing PD-L1 expression by the 22C3-IHC assay was considered as positive expression, and this was used as true positive results. The overall sensitivity and specificity of the SP142-IHC assay were 35% and 100%, respectively. When using 1% PD-L1-positive TCs as the cut-off point value, the SP142-IHC assay demonstrated a sensitivity of 57.8% and a specificity of 100%.

With the adoption of the SP142-IHC assay scoring algorithm, 10 (7.4%), 14 (10.4%), 37 (27.4%) and 74 (54.8%) of 135 samples evaluated using the SP142-IHC the assay were classified into groups that were strongly positive, moderately positive, weakly positive, and negative for PD-L1 expression, respectively. Only 69 (51.11%) of 135 samples produced the same results with the 22C3-IHC assay (κ value: 0.324, p < 0.001). In 60 (91.0%) of 66 discordant samples, PD-L1 status determined with the SP142-IHC assay was overestimated in the 22C3-IHC assay. In contrast, the PD-L1 status was underestimated by 22C3-IHC assay in six cases (McNemar-Bowker analysis: p = 0.001) (Table 3).

Association of PD-L1 Expression with Clinicopathological Characteristics. We determined the association of clinicopathological characteristics with PD-L1 expression based on each assay. PD-L1 expression based on the 22C3-IHC assay was significantly correlated with smoking status. However, the results from the SP142-IHC assay were different, indicating that PD-L1 expression was significantly correlated not only with smoking status but also with sex and histology (Table 4). Next, we evaluated the prognostic value of PD-L1 expression. No significant difference in survival was observed among patients with different PD-L1 expression by either assay (Fig. 4).
In this study, we use two clinically validated PD-L1 antibodies, 22C3 and SP142, and the corresponding IHC assay platforms to assess both the reproducibility/concordance and heterogeneity of PD-L1 protein expression. The results showed that the SP142-IHC assay was associated with significantly lower levels of PD-L1 detection in both TCs and ICs than was the 22C3-IHC assay in most of the cases, indicating that assays using the SP142 and 22C3 antibodies cannot replace each other for the stratification of patients for pembrolizumab or atezolizumab-based immune therapy. However, if the SP142-IHC assay results are positive and the 22C3-IHC assay results are negative, we can trust the interpretation.

Many studies have explored the topic regarding assay compatibility, but the conclusions have been inconsistent. One study demonstrated an analytical equivalence of TC membrane labeling between three validated, commercially available PD-L1 IHC diagnostic assays that use three different monoclonal antibodies, namely SP142-IHC assay, 22C3-IHC assay, and Total.

Table 3. PD-L1 IHC scoring according to the SP142-IHC assay scoring algorithm. 0 indicates negative, 1+ weakly positive, 2+ moderately positive, and 3+ strongly positive PD-L1 expression.

| Patient demographics (N=135) | 22C3 Negative N (%) | SP142 Positive N (%) | P-value | P-value |
|-----------------------------|---------------------|--------------------|--------|--------|
| Age (years)                 |                     |                    |        |        |
| >60 y                       | 15 (11.1)           | 37 (27.4)          | 0.247  | 0.000  |
| ≤60 y                       | 30 (22.2)           | 53 (39.3)          |        |        |
| Sex                         |                     |                    |        |        |
| Male                        | 36 (26.7)           | 62 (45.9)          | 0.221  | 0.003  |
| Female                      | 9 (6.7)             | 28 (20.7)          |        |        |
| Smoking status              |                     |                    |        |        |
| Never smoker                | 19 (14.7)           | 53 (41.1)          | 0.027  | 0.001  |
| Smoker                      | 26 (20.2)           | 31 (24)            |        |        |
| Histology                   |                     |                    |        |        |
| SCC                         | 24 (17.8)           | 29 (21.5)          | 0.053  | 0.001  |
| AD                          | 18 (13.3)           | 49 (36.3)          |        |        |
| Other subtypes              | 3 (2.2)             | 12 (8.9)           |        |        |
| TNM stage                   |                     |                    |        |        |
| I                           | 22 (16.3)           | 40 (29.6)          | 0.557  | 0.418  |
| II                          | 11 (8.1)            | 16 (11.9)          |        |        |
| III                         | 11 (8.1)            | 30 (22.2)          |        |        |
| IV                          | 1 (0.7)             | 4 (3.0)            |        |        |

Table 4. Association of Clinicopathological Features with PD-L1 Expression Based on the 22C3-IHC and SP142-IHC assays.

![Figure 4. Kaplan–Meier curves showing correlation between the OS of patients and PD-L1 expression determined by the 22C3-IHC and SP142-IHC assays.](image)

Discussion
In this study, we use two clinically validated PD-L1 antibodies, 22C3 and SP142, and the corresponding IHC assay platforms to assess both the reproducibility/concordance and heterogeneity of PD-L1 protein expression. The results showed that the SP142-IHC assay was associated with significantly lower levels of PD-L1 detection in both TCs and ICs than was the 22C3-IHC assay in most of the cases, indicating that assays using the SP142 and 22C3 antibodies cannot replace each other for the stratification of patients for pembrolizumab or atezolizumab-based immune therapy. However, if the SP142-IHC assay results are positive and the 22C3-IHC assay results are negative, we can trust the interpretation.

Many studies have explored the topic regarding assay compatibility, but the conclusions have been inconsistent. One study demonstrated an analytical equivalence of TC membrane labeling between three validated, commercially available PD-L1 IHC diagnostic assays that use three different monoclonal antibodies, namely,
Ventana SP263 with durvalumab, Dako 22C3 with pembrolizumab, and Dako 28-8 with nivolumab. Another study was designed to assess interobserver concordance and PD-L1 IHC staining patterns. Fifteen resected pulmonary carcinoma specimens were centrally stained with the 28-8, 22C3, SP142, and SP263 IHC assays according to clinical trial protocols. The data indicated that TCs can be reproducibly scored for PD-L1 expression based on IHC analysis in NSCLC tumors. No differences in interobserver concordance were observed among the examined assays, although the scoring of ICs had low concordance rates. However, that study did not include enough samples to be conclusive.

The Blueprint study was designed to evaluate the agreement of results from four PD-L1 IHC assays, namely, Dako 28-8, Dako 22C3, 29 Ventana SP263, and Ventana SP142. Thirty-eight non-interventional NSCLC cases were enrolled into the study, and each tumor was probed for PD-L1 using all four assays then scored by three pathologists. For TC staining, Dako 22C3, Dako 28-8 and Ventana SP263 assays demonstrated similar analytical performance, whereas Ventana SP142-IHC assay consistently labeled fewer TCs. Similar to our results, the SP142-IHC assay was associated with significantly lower levels of PD-L1 detection than was the 22C3-IHC assay. However, that study was different from ours, as the authors evaluated the percentage of ICs that were positively stained, while we evaluated positive staining in ICs based on the percentage of tumor area. The study performed by Rimm et al. to compare the performance of 4 PD-L1 assays (28-8, 22C3, SP142, and E1L3N) indicated that the assay using the SP142 antibody differed from the others, as it detected significantly lower PD-L1 expression in TCs and ICs. The limitation of the study is that the authors used the SP142 antibody on a Ventana platform not identical to the platform now approved by the FDA. In addition, the analysis was based on the average value as the gold standard to determine the sensitivity and specificity of various reagents. We used the standard for the 22C3-IHC as the gold standard with a 50% and 1% cut-off values and evaluated the sensitivity and specificity of PD-L1 labeling. We found that the sensitivity with the SP142-IHC assay is poor but that the specificity is better. Thus, if the SP142-IHC assay results are negative, the 22C3-IHC assay results will also be negative. Because the details of antibodies are proprietary information, we can only speculate that the poor concordance of the results with the two antibodies may be due to different antibody affinities, limited specificity, or distinct target epitopes.

It is important to note that we employed standardized platforms and reagents that are now approved by the FDA for this study and that this is currently the largest clinical study assessing the concordance between SP142-IHC and 22C3-IHC results in NSCLC.

The prognostic impact of PD-L1 protein expression and its association with clinicopathological features have not been conclusively elucidated. The reason for these discrepancies may be multifactorial. Differences in distributions of histologic subtypes, antibodies and cut-off values may account for some of these discrepancies. In our study, PD-L1 expression based on the 22C3-IHC assay was significantly correlated with smoking status. However, these results were different those obtained using the SP142-IHC assay, which indicated correlations of PD-L1 expression with sex, smoking status, and histology. Additionally, PD-L1 expression based on either IHC systems was not associated with patient prognosis. Our data represent the most direct evidence that demonstrates differences between two reagents and may partly explain the conflicting results regarding the correlation of PD-L1 expression and clinicopathological characteristics in other studies.

Our study has several limitations. First, all samples were obtained from patients who had not undergone PD1 or PD-L1 therapy. We did have access to therapeutic outcome data, and we could only evaluate the assay results by themselves without clinically meaningful comparisons of PD-L1 assays. Second, the vast majority of specimens were archived samples and not biopsies; these samples may not reflect the real situation in the clinic.

In conclusion, we identified that the use of different reagents for testing PD-L1 expression can be misleading; the two reagents cannot be interchangeably used with each other. Our findings also suggest that different reagents may account for inconsistent associations between the expression of PD-L1 and clinicopathological features.

References
1. Sul, J. et al. FDA Approval Summary: Pembrolizumab for the Treatment of Patients With Metastatic Non-Small-Cell Lung Cancer Whose Tumors Express Programmed-Death-Ligand 1. The Oncologist 21, 645–650 (2016).
2. Kumar, R. et al. Targeting the PD-1/PD-L1 axis in non-small cell lung cancer. Current problems in cancer 41, 111–124 (2017).
3. Aguiar, P. N. Jr., Santoro, I. L. & Tadokoro, H. et al. The role of PD-L1 expression as a predictive biomarker in advanced non-small-cell lung cancer: a network meta-analysis. Immunotherapy 8, 479–488 (2016).
4. Shukuya, T. & Carbone, D. P. Predictive Markers for the Efficacy of Anti-PD-1/PD-L1 Antibodies in Lung Cancer. Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer 11, 976–988 (2016).
5. Herbst, R. S., Baas, P. & Kim, D.-W. et al. Pembrolizumab versus docetaxel for previously treated, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. The Lancet Oncol 17, 643–650 (2016).
6. Reck, M., Rodriguez-Arue, D. & Robinzon, A. G. et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. The New England journal of medicine 375, 1823–1833 (2016).
7. Fehrenbacher, L., Spira, A. & Ballinger, M. et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. The Lancet 387, 1837–1846 (2016).
8. Cao, J. A., Boston, R. & Cane, P. et al. PD-L1 testing for lung cancer in the UK: recognizing the challenges for implementation. Histopathology 69, 177–186 (2016).
9. Taube, J. M., Klein, A. & Brahmer, J. R. et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature 515, 563–567 (2014).
10. Deterbeck, F. C., Chansky, K. & Groome, P. et al. The IASLC Lung Cancer Staging Project: Methodology and Validation Used in the Development of Proposals for Revision of the Stage Classification of NSCLC in the Forthcoming (Eighth) Edition of the TNM Classification of Lung Cancer. Journal of thoracic oncology 11, 1433–1446 (2016).
11. Travis, W. D., Brambilla, E. & Nicholson, A. G. et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Clinical, Genetic and Radiologic Advances Since the 2004 Classification. Journal of thoracic oncology 10, 1243–1260 (2015).
12. Scher, A. H., Dietel, M. & Heukamp, L. C. et al. Harmonized PD-L1 immunohistochemistry for pulmonary squamous-cell and adenocarcinomas. Mod Pathol 29, 1165–1172 (2016).
13. Hirsch, F. R., McElhany, A. & Sanftor., D. et al. PD-L1 Immunohistochchemistry Assays for Lung Cancer: Results from Phase 1 of the Blueprint PD-L1 IHC Assay Comparison Project. J Thorac Oncol 12, 208–222 (2017).
14. Rimm, D. L., Han, G. & Taube, J. M. et al. A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. *JAMA Oncol* 3, 1051–1058 (2017).
15. Azuma, K., Ota, K. & Kawahara, A. et al. Association of PD-L1 overexpression with activating EGFR mutations in surgically resected nonsmall-cell lung cancer. *Annals of oncology: official journal of the European Society for. Medical Oncology* 25, 1935–1940 (2014).
16. Chen, N., Fang, W. & Zhan, J. et al. Upregulation of PD-L1 by EGFR Activation Mediates the Immune Escape in EGFR-Driven NSCLC: Implication for Optional Immune Targeted Therapy for NSCLC Patients with EGFR Mutation. *Journal of Thoracic Oncology* 10, 910–923 (2015).
17. Li, J. et al. Identification of high-quality cancer prognostic markers and metastasis network modules. *Nature communications* 1, 34 (2010).
18. Kerr, K. M. et al. Programmed Death-Ligand 1 Immunohistochemistry in Lung Cancer: In what state is this art? *Journal of thoracic oncology* 10, 985–989 (2015).

**Acknowledgements**
This work was supported by the Natural Science Foundation of China (grant number 81372599), Fujian Provincial health and family research talent training program (grant number 2015-ZQN-ZD-9), and the Natural Science Foundation of Fujian Province (grant number 2015J01434).

**Author Contributions**
Haipeng Xu, Gen Lin, Chao Li, and Cheng Huang wrote the main text and Weifeng Zhu, Xiandong Lin, and Dan Hu performed immunohistochemistry. Qian Miao Xirong Fan and Biao Wu collected cases. All authors reviewed the manuscript.

**Additional Information**
**Competing Interests:** The authors declare that they have no competing interests.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit [http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).

© The Author(s) 2017