

2017

The "eyes" have it?-intra- and inter-observer reproducibility of the PD-L1 companion diagnostic assay

Steven G. Gray

Mark Wright

Stephen Finn

Follow this and additional works at: <https://arrow.tudublin.ie/scschbioart>



Part of the [Oncology Commons](#)

This Article is brought to you for free and open access by the School of Biological Sciences at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)



The “eyes” have it?—intra- and inter-observer reproducibility of the PD-L1 companion diagnostic assay

Steven G. Gray^{1,2,3,4}, Mark C. Wright⁵, Stephen P. Finn^{1,6,7}

¹Thoracic Oncology Research Group, Rm 2.09, Trinity Translational Medical Institute, Trinity Centre for Health Sciences, St. James's Hospital, Dublin D08 W9RT, Ireland; ²HOPE Directorate, St. James's Hospital, Dublin D08 RT2X, Ireland; ³Department of Clinical Medicine, School of Medicine, Trinity College Dublin, Dublin D02 R590, Ireland; ⁴School of Biological Sciences, Dublin Institute of Technology, Dublin D08 NF82, Ireland; ⁵Histopathology, Central Pathology Laboratory, ⁶Cancer Molecular Diagnostics Laboratory, St. James's Hospital, Dublin D08 RX0X, Ireland; ⁷Department of Histopathology & Morbid Anatomy, Trinity College Dublin, Dublin D08 RX0X, Ireland

Correspondence to: Steven Gray. Thoracic Oncology Research Group, Rm 2.09, Trinity Translational Medical Institute, Trinity Centre for Health Sciences, St. James's Hospital, Dublin D08 W9RT, Ireland. Email: sgray@stjames.ie; Stephen Finn. Cancer Molecular Diagnostics Laboratory, Central Pathology Laboratory (CPL) Building, St. James's Hospital, Dublin D08 RX0X, Ireland. Email: Stephen.Finn@tcd.ie.

Comment on: Cooper WA, Russell PA, Cherian M, *et al.* Intra- and Interobserver Reproducibility Assessment of PD-L1 Biomarker in Non-Small Cell Lung Cancer. *Clin Cancer Res* 2017;23:4569-77.

Submitted Jun 23, 2017. Accepted for publication Jul 04, 2017.

doi: 10.21037/tcr.2017.07.06

View this article at: <http://dx.doi.org/10.21037/tcr.2017.07.06>

The role of the immune system in cancer is well established. Indeed, most potential cancer causing cells are detected and removed from our bodies by our immune system in a process called “immune surveillance”. However, at some point, tumors manage to evade our immune system often by expressing signals that inhibit the anti-tumor immune response (1). “The scientific turning point for cancer immunotherapy came with the understanding that T cell immune responses are controlled through on and off switches, so-called ‘immune checkpoints’ that protect the body from possibly damaging immune responses” (2). Blockade of these checkpoints has emerged as a new paradigm for the treatment of a cancer, including NSCLC (3).

One of the most exciting therapeutic developments currently in NSCLC involves targeting the checkpoint involving the PD-1 [programmed death 1 (PD-1) receptor] protein and its ligand, programmed death ligand 1 (PD-L1). Monoclonal antibodies to either PD-1 or PD-L1 have shown impressive response rates in NSCLC patients, and have recently received regulatory approval to treat NSCLC in both the EU and US (4,5).

Testing for expression of these checkpoint inhibitor targets has proceeded apace with either companion assays or complementary assays. A companion assay is a necessary requirement for use of the corresponding drug, whilst complementary assays are recommended in order

to optimize appropriate patient selection, but are not mandatory (6).

A positive PD-L1 IHC test has been shown to be predictive of better responses and in many cases better patient outcome for anti-PD1 and anti-PD-L1 based therapies (7). In October 2015, a PD-L1 IHC test was approved by the FDA as a companion diagnostic for Pembrolizumab in treating advanced NSCLC (PD-L1 IHC 22C3 pharmDx™) (8).

However, accurate measurement and scoring of PD-L1 protein expression are plagued by various technical and biological pitfalls (7,9,10). Given the projected economic costs for checkpoint inhibitors in NSCLC are of the order of \$130,511 for Pembrolizumab (11), this poses significant challenges for hospital laboratories and pathologists, only one of which is the issue of intra- and inter-observer reproducibility of scoring for such companion diagnostics.

In an article to be published in *Clinical Cancer Research (In Press)*, Cooper and colleagues (12), assessed the FDA approved PD-L1 companion diagnostic (PD-L1 IHC 22C3 pharmDx™). Ten pathologists examined two NSCLC samples comprising two sample sets for the two established cut-points for positivity (1% and 50%, with n=60 samples for each) for both inter- and intra- observer reproducibility, and further tested whether a one-hour training session could affect assessment. Scoring for this study was only based on a “tumor proportion score (TPS)”. Following analysis,

for the 1% cut-off sample set the authors found that intra-observer reproducibility had an overall percent agreement of 89.7% (95% CI: 85.7–92.6), whilst that for the 50% cut-off was 91.3% (95% CI: 87.6–94.0). For inter-observer reproducibility the values for overall percent agreement at the 1% cut-off were 84.2% (95% CI: 82.8–85.5), and at 50% were 81.9% (95% CI: 80.4–83.3) (12). When compared against a “gold standard” PD-L1 TPS, the concordance for the 1% sample set was 84.3% for sensitivity (95% CI: 80.2–88.5) and 91.3% for specificity (95% CI: 88.2–94.5). For the 50% sample set the concordance for sensitivity was 56.3% (95% CI: 50.7–62.0) and specificity was 94% (95% CI: 91.3–96.7). Surprisingly, the impact of a 1-hour training session taken prior to a second assessment of the samples was found to have no impact on overall percent agreements for the 1% cut-off sample set, and only slight improvement for the 50% cut-off set (rising from 78.3% to 81.7%). When training was assessed against the gold standard, minimal effects were observed (1–87.3% as opposed to 87.7% pre versus post training; 50–75.3% versus 78.7%) (12).

How does this compare against other studies?

The results from the various clinical trials of Pembrolizumab have suggested that sensitivity of the 22c3 assay is 76% and the specificity is 60% (6). Most recently, Rimm and colleagues also examined the PD-L1 IHC 22C3 pharmDx™ assay as part of a multi-institutional, pathologist assessment of four of the current IHC assays for PD-L1 expression in a cohort of n=90 NSCLC specimens (13). In this study 13 pathologists also scored sections according to TPS. The concordance between pathologist's scores for tumor cells for the 22c3 assay was 0.822 (95% CI: 0.873–0.891). To estimate sensitivity, the authors “defined the median pathologist's score as “truth” and calculated the correctly predicted proportion of positive cases as an analogue for sensitivity and a correctly predicted proportion negative as an analogue for specificity” (13). By this method they found that the assay had 90% to 95% sensitivity for either the 1% or 50% cut-off. When specificity was examined the 1% cut-off had 70–80% specificity, while the 50% cut-off had greater than 95% specificity (13). Another attempt to compare different PD-L1 IHC is the Blueprint project (14). In this study this study three pathologists assessed TPS in (n=39) NSCLC cases and used the 1% cut-off for analysis of concordance. In this regard, the 22c3 assay achieved 100% concordance (14).

There are limitations to all of the studies described above, which necessitates caution with respect to direct comparisons between these studies. For example, the

Blueprint study and the study by Rimm and colleagues were attempts to address the issue of concordance between four separate PD-L1 assays (13,14), whereas in contrast the study by Cooper and colleagues had a specific focus on one particular assay. In addition it must be noted that the Blueprint study as published was considered to be a feasibility study, and was therefore not powered for statistical analysis.

One study used exclusively full-face sections from surgically resected cancers (13), one used a mixture of full-face sections and biopsy material (14), whilst the remaining study utilized a tissue microarray (12). One well established tenet of PD-L1 staining is that both inter- and intra-tumoral heterogeneity of staining is common (9). In this regard, the number of replicate cores per patient in the TMA used by Cooper *et al.*, is not stated, and if only one core per patient was used may have added a bias to the analysis (12). Indeed assessment of cores rather than full face sections may represent a relatively ‘easier task’ for the participating pathologists given the smaller area of tumor needed to derive the denominator of the TPS.

One issue between the study of Cooper *et al.* (12), and Rimm *et al.* (13), was that the intra-assay variation between pathologists for the percentage of tumor staining was better for the 50% cut-point versus the 1% cut-point in Rimm *et al.* (13,15), whilst in the study by Cooper and colleagues, pathologists mostly underscored the samples in the 50% cut-point range (12).

Another limitation to making direct comparisons between these two studies potentially relates to the statistical methodology used. Cooper *et al.*, used Cohen's kappa coefficient, whilst due to the nature of the analysis/cross comparisons Rimm *et al.*, used the Fleiss kappa coefficient and Kendall's concordance coefficient (12,13).

In the initial development of the PD-L1 IHC 22C3 pharmDx™ assay, the in-house analytical validation comprising of inter-instrument, inter-operator, inter-day, inter-lot, intra-day and intra-run variations found overall percent agreement was 100%. When reproducibility was tested at different external laboratories, the concordance was slightly lower with an overall percent agreement of 88.8% (8). The results from the three additional studies appear to support these initial observations appear to indicate that the PD-L1 IHC 22C3 pharmDx™ assay is robust with good reproducibility at both 1% and 50% cut-points. However, it must be noted that all of the pathologists taking part in these studies have significant experience in scoring IHC biomarkers. In effect, once the pathologists “eye” is trained in, the concordance and reliability of assays such as the PD-L1 IHC 22C3 pharmDx™ assay are robust.

Therefore, should all assays for PD-L1 testing be conducted by experienced personnel, or should there be centralized testing for patients for suitability for Pembrolizumab treatment? Finally, the question still remains, what to do for the approximately 15% of patients, classified as PD-L1 negative by IHC, that actually respond to therapy? (7).

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned and reviewed by the Section Editor Shaohua Cui (Department of Pulmonary Medicine, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr.2017.07.06>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Mahoney KM, Rennert PD, Freeman GJ. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discov* 2015;14:561-84.
- Hoos A. Development of immuno-oncology drugs—from CTLA4 to PD1 to the next generations. *Nat Rev Drug Discov* 2016;15:235-47.
- Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* 2015;27:450-61.
- Kazandjian D, Suzman DL, Blumenthal G, et al. FDA Approval Summary: Nivolumab for the Treatment of Metastatic Non-Small Cell Lung Cancer With Progression On or After Platinum-Based Chemotherapy. *Oncologist* 2016;21:634-42.
- Sul J, Blumenthal GM, Jiang X, 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. *Oncologist* 2016;21:643-50.
- Diggs LP, Hsueh EC. Utility of PD-L1 immunohistochemistry assays for predicting PD-1/PD-L1 inhibitor response. *Biomark Res* 2017;5:12.
- Kerr KM, Hirsch FR. Programmed Death Ligand-1 Immunohistochemistry: Friend or Foe? *Arch Pathol Lab Med* 2016;140:326-31.
- Roach C, Zhang N, Corigliano E, et al. Development of a Companion Diagnostic PD-L1 Immunohistochemistry Assay for Pembrolizumab Therapy in Non-Small-cell Lung Cancer. *Appl Immunohistochem Mol Morphol* 2016;24:392-7.
- Ilie M, Hofman V, Dietel M, et al. Assessment of the PD-L1 status by immunohistochemistry: challenges and perspectives for therapeutic strategies in lung cancer patients. *Virchows Arch* 2016;468:511-25.
- Villaruz LC, Socinski MA. The clinical utility of PD-L1 testing in selecting non-small cell lung cancer patients for PD1/PD-L1-directed therapy. *Clin Pharmacol Ther* 2016;100:212-4.
- Tartari F, Santoni M, Burattini L, et al. Economic sustainability of anti-PD-1 agents nivolumab and pembrolizumab in cancer patients: Recent insights and future challenges. *Cancer Treat Rev* 2016;48:20-4.
- Cooper WA, Russell PA, Cherian M, et al. Intra- and Interobserver Reproducibility Assessment of PD-L1 Biomarker in Non-Small Cell Lung Cancer. *Clin Cancer Res* 2017. [Epub ahead of print].
- Rimm DL, Han G, Taube JM, 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* 2017. [Epub ahead of print].
- Hirsch FR, McElhinny A, Stanforth D, et al. PD-L1 Immunohistochemistry Assays for Lung Cancer: Results from Phase 1 of the Blueprint PD-L1 IHC Assay Comparison Project. *J Thorac Oncol* 2017;12:208-22.
- Sica GL, Ramalingam SS. Assays for PD-L1 Expression: Do All Roads Lead to Rome? *JAMA Oncol* 2017;3:1058-59.

Cite this article as: Gray SG, Wright MC, Finn SP. The “eyes” have it?—intra- and inter-observer reproducibility of the PD-L1 companion diagnostic assay. *Transl Cancer Res* 2017;6(Suppl 6):S1027-S1029. doi: 10.21037/tcr.2017.07.06