The Challenges of Implementing a PD-L1 Proficiency Testing Program in Australia

Julia Pagliuso, [1, a, @] Suzanne Parry, [2] Zenobia Haffajee, [1] Tony Badrick, [1] Keith Miller, [2]

@ corresponding author, & equal contributor

Author information
1. The Royal College of Pathologists of Australasia Quality Assurance Programs (RCPAQAP) - The Royal College of Pathologists of Australasia Quality; St Leonards, NSW 2065, Australia
2. UK NEQAS Immunocytochemistry (ICC) and In-Situ Hybridisation (ISH) - Finsbury Business Centre; London, United Kingdom

[a] julia.pagliuso@rcpaqap.com.au

Abstract

Background . One important initiative that commenced at the Royal College of Pathologists of Australasia Quality Assurance Program (RCPAQAP) in 2017 was the collaboration with United Kingdom National External Quality Assessment Scheme (UK NEQAS) Immunocytochemistry (ICC) and In-Situ Hybridization (ISH) for the challenging implementation of a PD-L1 immunohistochemistry (IHC) proficiency testing program for non-small cell lung carcinoma (NSCLC). A RCPAQAP participant survey in 2016 showed that only eight laboratories were performing PD-L1 testing. The aim of the collaboration was to increase the sample size of the pilot program to provide meaningful results that could be reported back to RCPAQAP participants with appropriate recommendations. Other challenges of assessment included standardising the clinical cut-offs for positivity for each commercial assay, interpretation of laboratory developed tests (LDTs), using appropriate tissue to cover the critical interpretation points for each assay, interchangeability of clones and interpretation proficiency testing.

Methods . The use of a 'Gold Standard' for each commercial assay was used as a baseline to compare participant results and tumour proportion score bin categories were implemented to harmonise interpretation across clones.

Conclusions . The findings of the pre-pilot test suggest that the use of a clinically validated PD-L1 IHC assay performs better during assessment than adopting a laboratory developed test (LDT). The assessment committee also concluded that tonsil showed a better dynamic range of positivity than placenta. It was acknowledged that participants are limited by the platforms they have available and so it was suggested that validating an optimal method against the clinical assay and continual verification of the test may produce the expected result. The next big challenge is to extend proficiency testing from technical to interpretation. This is being implemented globally via the International Quality Network for Pathology (IQNPath) with participation through local External Quality Assurance programs, including RCPAQAP.

Keywords

lung — NSCLC — PD-L1 — Immunohistochemistry — quality — RCPAQAP

Introduction

Lung cancer is the most common cancer in the world and represents the most common cause of death from cancer worldwide [1]. Approximately 85% of all lung cancer cases are non-small cell type (NSCLC) and traditionally the treatment of this category of lung cancer was limited to radiotherapy, chemotherapy, or a combination of both [2]. Although much progress has been recently made for lung cancer such as molecularly targeted

therapies, patients with lung cancer are still facing a relatively low 5-year survival rate at 17.4% [3,4]. Recent approaches to NSCLC management has focused on targeting immune checkpoint inhibitors.

Programmed death 1 (PD-1), a member of the CD28 family, is a key immune checkpoint receptor expressing on the surface of the activate T, B and natural killer (NK) cells and plays a crucial role in tumour immune escape [3].

Programmed death ligand 1 (PD-L1), is upregulated in different types of tumours, including NSCLC [3,5]. PD-L1 binds to PD-1 to reduce the immune response by inducing T-cell apoptosis or exhaustion [3]. Efforts to use monoclonal antibodies (mAbs) to target and block these immunoinhibitory interactions have led to a new era of immunotherapy-based agents for cancer therapy [6,7]. Current data show that patient outcomes are generally better with these therapies when there is an increase in PD-L1 expression as measured by IHC [2].

Challenge 1: Multiple PD-L1 Biomarkers for multiple therapies. PD-L1 is unique to other biomarkers in that at least four different therapies for NSCLC have been developed or are in the development phase targeting PD-1/PD-L1, and have been clinically validated with four different companion or complementary PD-L1 immunohistochemistry (IHC) assays to determine patient eligibility and likelihood of response to their respective therapies [8]. Any of the fully human anti-PD1 mAb BMS-936558 (Nivolumab), the humanised anti-PD-1 antibody MK-3945 (Pembrolizumab) and anti-PD-L1 mAbs Atezolizumab and Durvalumab [6] immune checkpoint inhibitor drugs could potentially be applied as patient treatment by oncologists. Although Atezolizumab is not commercially available, IHC interpretation is complicated by the fact that different clones of mAbs raised against the same protein will be specific for different protein epitopes.

Consequently, one PD-L1 IHC test may not necessarily perform in the same way as another [2]. Biomarker studies conducted in the trials of Nivolumab used the anti-PDL1 IHC antibody clone 28-8 (Dako, Glostrup, Denmark). Alternatively, Pembrolizumab studies used a different anti-PD-L1 Dako clone, 22C3. Durvalumab and Atezolizumab had complementary diagnostic tests based on different clones of anti-PD-L1 – Ventana SP263 and SP142 (Tucson, Arizona), respectively [2].

Challenge 2: Different scoring systems. These clones use different scoring systems and have different cut-off thresholds for defining positivity for the application of each drug. In Nivolumab trials, tumour cell staining for PD-L1 was assessed using different thresholds (≥1%, ≥5% and ≥10%) to define positive staining [2,9]. Pembrolizumab trials considered two 'positive' thresholds of tumour cell staining (≥1% and ≥50%) and the published data support the use of a threshold of 50% or greater for clinical use [2,10]. Alternatively, the positive threshold for Durvalumab was defined as tumour cell staining of 25% or greater [11] and Atezolizumab is even more detailed with an assessment of both tumour cells and/or tumour-associated immune cells required using the SP142 clone. For tumour cells, four different grades of staining have been considered in clinical trials, defined around cut points of 1%, 5% and 50% tumour cell staining. Immune cell staining is defined with cut-offs at 1%, 5% and 10% [2,12]. The possibility of a choice of four different drugs and four different approaches to PD-L1 IHC testing would bring significant challenges to oncologists, laboratories performing IHC testing and pathologists interpreting the complementary or companion diagnostic test. To add to the complications surrounding PD-L1 IHC testing, many laboratories may have been employing LDTs if the assay-specific platform was not available or to save money on testing. Consequently, it became important for external quality assurance (EQA) of pathology providers to ensure that both technical and interpretation aspects of this testing were being performed satisfactorily and to devise a method to incorporate the complications surrounding the variables of testing within the program.

One important initiative that commenced at the Royal College of Pathologists of Australasia Quality Assurance Program (RCPAQAP) in 2017 was the collaboration with United Kingdom National External Quality Assessment Scheme (UK NEQAS) Immunocytochemistry (ICC) and In-Situ Hybridization (ISH) for this challenging implementation of a PD-L1 IHC proficiency testing program for NSCLC.

Challenge 3: Participation. An RCPAQAP participant survey in 2016 showed that only eight laboratories were performing PD-L1 testing. It would not be viable or meaningful to establish an external quality assurance (EQA) program for only eight participants with the possibility of as many variations in method submissions. UK NEQAS were also at the beginnings of establishing their own program and so the aim of the collaboration with UK NEQAS was to increase the sample size of the pre-pilot program to provide meaningful results and establish a set of guidelines to help harmonise the assessment process. Thirteen Australian laboratories were included in the UK NEQAS pre-pilot in early 2017, which attracted a total of 47 participants. This increased number of combined participants allowed for a meaningful comparison of results between submissions.

Methods
Pre-pilot participants were provided unstained

formalin-fixed, paraffin-embedded (FFPE) tissue sections from two different multi-blocks (1 and 2).

The block was a combination of cell lines, tonsil and NSCLC tissue (A-H) as shown in Figure 1.

Participants were also asked to submit their methodology with the returned stained slides. The assessment panel included sixteen expert pathologists and scientists.

The pre-assessment meeting included discussions on how to approach the various complications of PD-L1. Outcomes of the discussion included: (a) scoring each individual core/section based on the tumour proportion score (TPS) regardless of intensity and (b) immune cells would only be counted when assessing the SP142 assay, but most importantly, (c) a method to harmonise the clinical cut-offs for positivity was established.
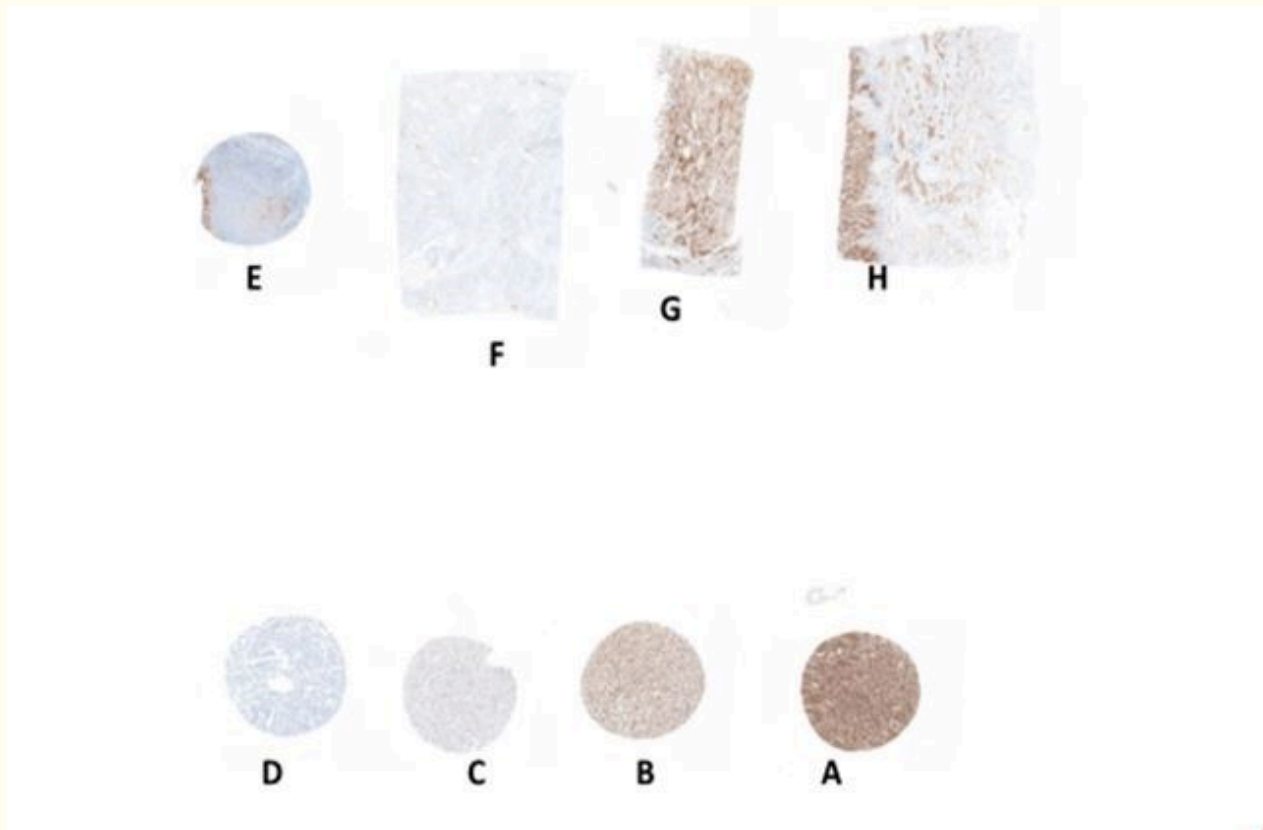

Figure 1

Figure 1 caption

Example of PD-L1 expression for multiblock 2

Challenge 4: Varying clinical cut-offs for positivity. Rather than assessing each clone according to the various clinical cut-offs, a harmonised approach was established which could be used during the assessment of all clones. The TPS was categorised using a series of bins (Table 1) which were set at a TPS range that allowed for assessment of specificity of the result submitted. The TPS was then applied to each core/section on the gold standard for each commercial assay (Figure 2).
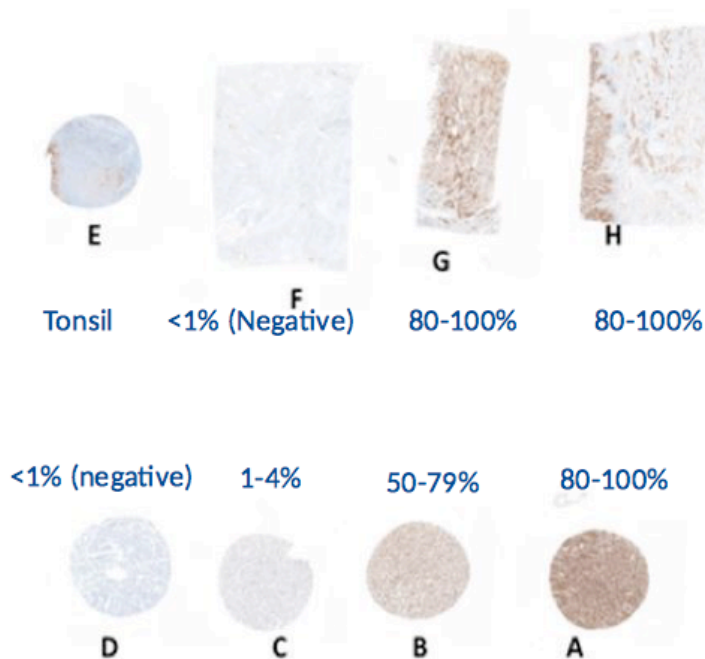
Figure 2

Figure 2 caption

TPS bin category applied to Gold sections

Table 1

TUMOUR PROPORTION SCORE (TPS) BINS	IMMUNE CELL (IC) SCORE BINS
<1% (negative)	<1% (negative)
1-4%	1-4%
5-9%	5-9%
10-24%	≥10%
25-49%	
50-79%	
80-100%	

Challenge 5: Applying a baseline comparator. To create the gold standards, each block and at every 25th serial level, sections were stained by the manufacturers of the Dako/Agilent 22C3 and 28-8 and the Ventana/Roche SP263 and SP142 approved PD-L1 assays. These 'Golds' were then used as a baseline to compare participant results.

Assessment

The pre-pilot assessment consisted of two groups of assessors, each consisting of at least one PD-L1 specialist pathologist trained in interpretation of PD-L1 assays. Each section/core was assessed on:

(a) Bin category for each test core/section matching the corresponding gold bin category and (b) technical quality. Opinions were given and a consensus score out of 5 was provided against the scoring criteria described in Table 2. The UK NEQAS distributed tonsil control tissue (sample E) was assessed as either acceptable, borderline or unacceptable.

Challenge 6: Laboratory Developed Tests (LDTs). It was expected that submissions would include LDTs as they are generally less expensive than the commercial assay. The challenge during the assessment was that there is no standardisation or clear gold standard comparator for LDTs (in-house in vitro diagnostic medical devices-IVDs), apart from the commercial kits themselves. Proficiency testing becomes extremely important in this scenario to provide oversight and promote high quality and consistent PD-L1 IHC results across antibodies and test platforms and in a variety of settings. The gold standard slides for each of the commercial assays were also used to compare LDT results.

Table 2

SCORE	PD-L1 DEMONSTRATION
4-5	Good/Excellent demonstration of PD-L1
3	Acceptable demonstration – slightly weak/strong staining; some of the required components may be missing of there may be non-specific/inappropriate staining present.
1-2	Failure to demonstrate the required PD-L1 components

Results

A breakdown of pass rates and methodologies are summarised in Table 3 and Figure 3. The shaded cells in Table 3 represent the commercial assays and the white cells represent the various LDT methods submitted for assessment. Twenty out of 47 participants submitted an LDT stained slide for the pre-pilot.

Table 3

PD-L1 Assay	Automation	Detection Kit	Good/ Excellent	Acceptable/ Borderline	Unacceptable	n=47
Dako/Agilent 22C3 PharmDx Assay	Dako Autostainer Link 48	Dako Envision FLEX+	7 (78%)	2 (22%)		n=9
Dako/Agilent 22C3 mAB concentrate	Dako Autostainer Link 48	Dako Envision	-	1 (50%)	1 (50%)	n=2
	Leica BOND-MAX	Leica Bond Polymer Refine	-	-	1 (100%)	n=1
	Ventana Benchmark Ultra/XT	Ventana Optiview	1 (14%)	3 (43%)	3 (43%)	n=7
	Manual Stain	Dako REAL envision	-	1 (50%)	1 (50%)	n=2
Dako/Agilent 22-8 PharmDx Assay	Dako Autostainer Link 48	Dako Envision FLEX+		1 (100%)		n=1
Ventana/Roche SP263 Assay	Ventana Benchmark	Ventana Optiview	8 (57%)	4 (29%)	2 (14%)	n=14
Ventana/Roche SP142 Assay	Ventana Benchmark	Ventana Optiview	3 (100%)	-	-	n=3
Spring Bioscience SP142 mAb Concentrate	Ventana Benchmark Ultra	Ventana Optiview	-	-	1 (100%)	n=1
	Ventana Benchmark XT	Ventana Ultraview	-	-	1 (100%)	n=1
Abcam 28-8 mAb Concentrate	Ventana Benchmark	Ventana Ultraview	-	-	1 (100%)	n=1

PD-L1 Assay	Automation	Detection Kit	Good/Excellent	Acceptable/Borderline	Unacceptable	n=47
	XT					
28-8 Supplier not specified	Not specified	Not specified		1 (100%)		n=1
Biocare CAL10 mAb Concentrate	Ventana Benchmark Ultra	Ventana Ultraview	-	1 (100%)	-	n=1
	Leica Bond III	Leica Bond Polymer Refine	-	1 (100%)	-	n=1
Cell Signaling Technologies mAb E1L3N Concentrate	Ventana Benchmark Ultra	Ventana Ultraview			1 (100%)	n=1
	Leica Bond III	Leica Bond Polymer Refine	1 (100%)	-	-	n=1

The two most common approved assays used were Dako 22C3 and Ventana SP263 which performed well with only zero out of nine and two out of fourteen participants obtaining unacceptable results respectively. When Dako 22C3 was used as an LDT, six out of twelve participants received an unacceptable result. From the other clones employed, only one (E1L3N used on the Leica Bond III platform) achieved a good/excellent result. Three out of thirteen Australian participants received a good/excellent result and seven out of thirteen results submitted from Australia were LDTs. Overall, the pass rates show that participants using the PD-L1 approved assays achieved higher results than laboratories using LDT methods. Figures 4-6 (4A, 4B, 5A, 5B, 6A, 6B) illustrate participant submissions compared with the gold standard. It can be seen that poor technical quality could force an incorrect interpretation of the TPS which may impact treatment protocols. This is particularly evident in Figure 6 (6A, 6B) where the patient may not have been offered first-line therapy with Pembrolizumab. Generally, it was noted that tonsil was a preferred control over placenta to show varying levels of PD-L1. Acceptable tonsil staining should show moderate-to-strong PD-L1 staining in crypt epithelial cells and diffuse staining in the germinal centres. The pre-pilot was followed by a pilot (Run 119) in the latter part of 2017.

Challenge 7: Donation of tissue. IHC EQA providers rely on the donation of tissues from participants and advisory committees. Due to the lack of tissue, only the five Australian participants with an unsatisfactory assessment in the pre-pilot were invited to participate in the pilot plus two new laboratories who are regular donors to the RCPAQAP programs. A general improvement was seen between the pre-pilot and pilot results, with the exception of the E1L3N clone (Figure 7).

Challenge 8: Interchangeability of clones. The Blueprint Phase 2 project [13], verified that Dako 28-8/22C3 and Ventana SP263 show very similar levels of PD-L1 expression on tumour cells suggesting the interchangeability of these three assays. In contradiction, UK NEQAS found that in run 119, differences in PD-L1 expression were seen for cell lines F and G when applying the TPS to the gold standard slides. This was an unexpected challenge and added a new dimension to the assessment. In this scenario, it became extremely important that participants submit the appropriate methodology with the slide submission to be assessed against the correct assay.

The Australian pre-pilot and pilot results are compared in Table 4. Two out of the three participants who indicated no change to their protocol for the pilot continued to receive a score of 2 for the pilot. One participant who changed from the 22C3 LDT method to the SP263 commercial assay achieved an improvement in score from 2 to 4. One participant switched from a 22C3 LDT in the pre-pilot to an SP263 LDT in the pilot and did not show an improvement in score. It is noted that other EQA schemes have also shown that the commercial assays generally performed better than LDTs, but also that improvement is evident between surveys. Nordic Immunohistochemistry Quality Control (NordiQC) results in Run C1 2017 [14] showed a pass rate of 80% for approved assays and 20% for LDTs. In Run C2 2018 [15], the pass rate improved in both categories showing a pass rate of 95% for approved assays and 73% for LDTs. NordiQC also showed a preference for tonsil as a control over placenta.

Table 4

PRE-PILOT				PILOT			
PD-L1 Assay	Automation	Detection Kit	Pre-Pilot score	PD-P1 Assay	Automation	Detection Kit	Pilot score
Dako/Agilent **22C3** mAB Concentrate	Dako Autostainer Link 48	Dako Envision	2	No change			2
Dako/Agilent **22C3** mAB Concentrate	Leica Bond Max	Leica Bond Polymer Refine	2	Ventana/ Roche **SP263** Assay	Leica BondMax	Leica Bond Polymer Refine	2
Dako/Agilent **22C3** mAB Concentrate	Ventana Benchmark Ultra/XT	Ventana Optiview	2	Ventana/ Roche **SP263** Assay	Ventana Benchmark	Ventana Optiview	4
Ventana/Roche **SP263** Assay	Ventana Benchmark	Ventana Optiview	2	No change			2
Ventana/Roche **SP263** Assay	Ventana Benchmark	Ventana Optiview	2	No change			4

Challenge 9: Interpretation proficiency testing It should be noted that participants' interpretation of the TPS was not assessed since UK NEQAS ICC and ISH are purely a technical EQA scheme. Like many other EQA program providers, RCPAQAP is a member of International Quality Network for Pathology (IQNPath), which is an international multi-stakeholder expert group focused on improving the quality of clinical biomarker testing. Amongst other ventures, IQNPath is creating a digital, educational self-assessment for pathologists to test TPS interpretation for the four FDA approved PD-L1 assays [16]. A pilot for this educational portal is expected in 2018.
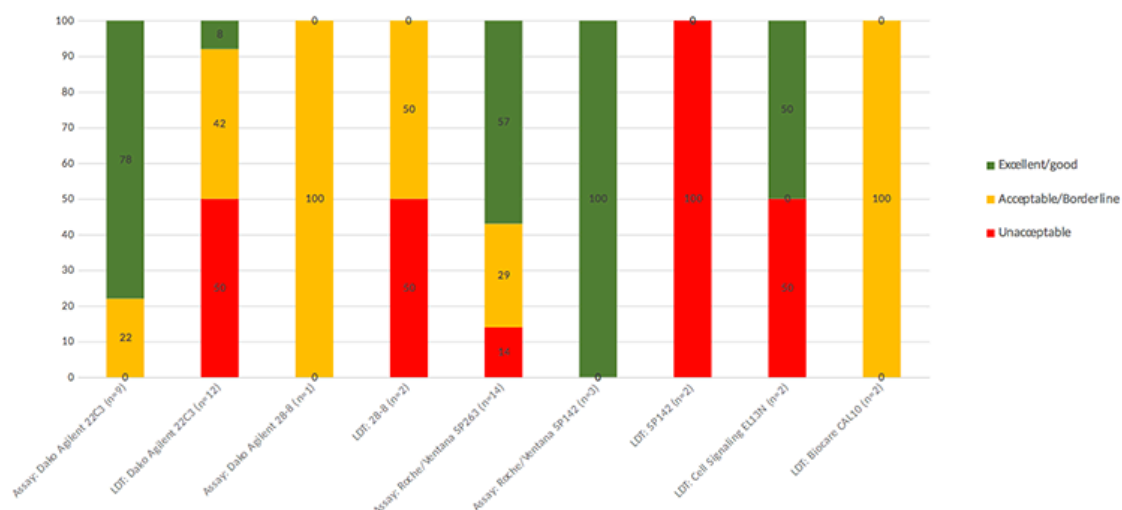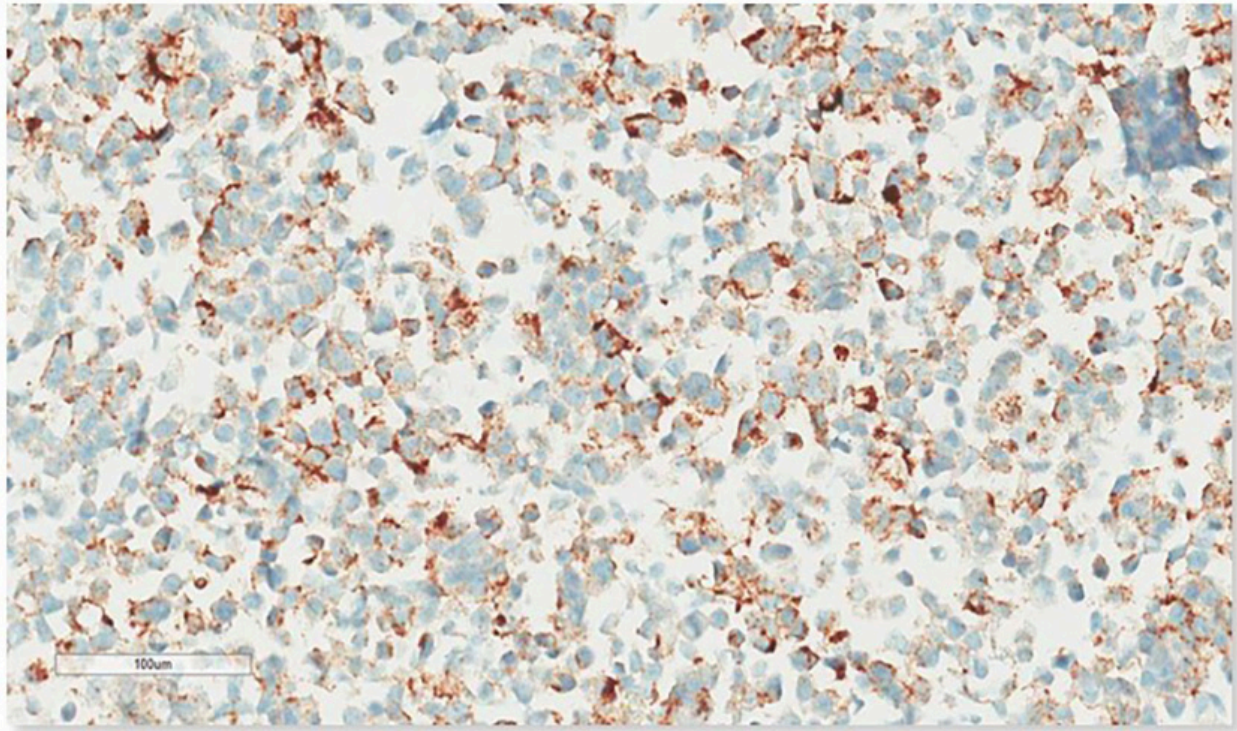

Figure 3

Figure 4A
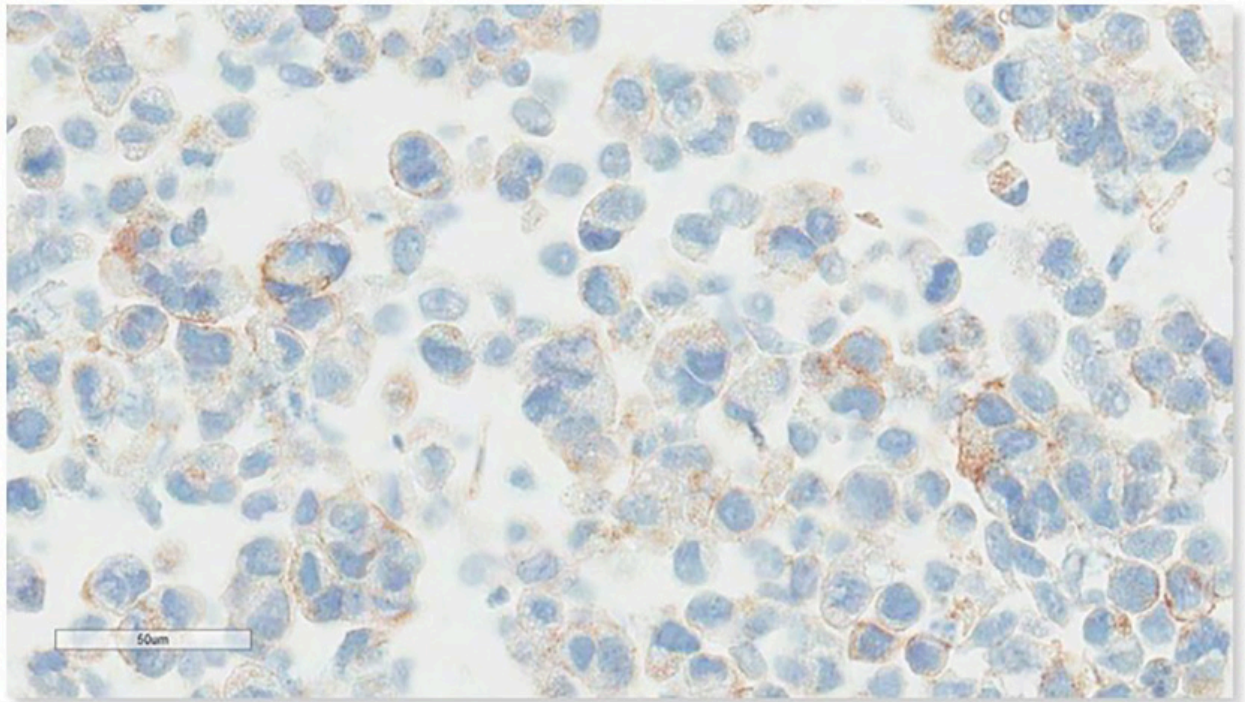
Figure 4B

Figure 4B caption

Participant submission 22C3 LDT 'C'. Over antigen retrieved. (Mark 2/5)
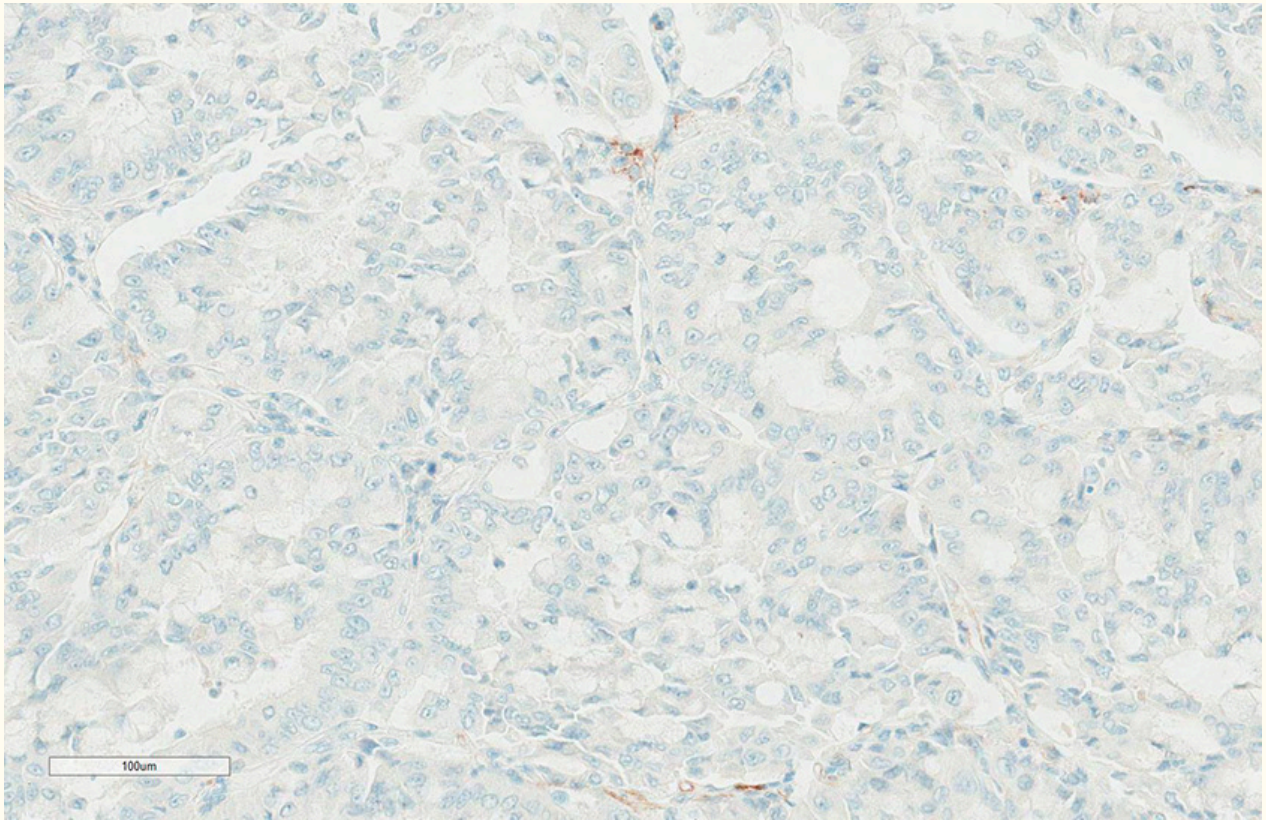
Figure 5A

Figure 5A caption

Ventana/Roche SP263 Gold 'F' <1% TPS
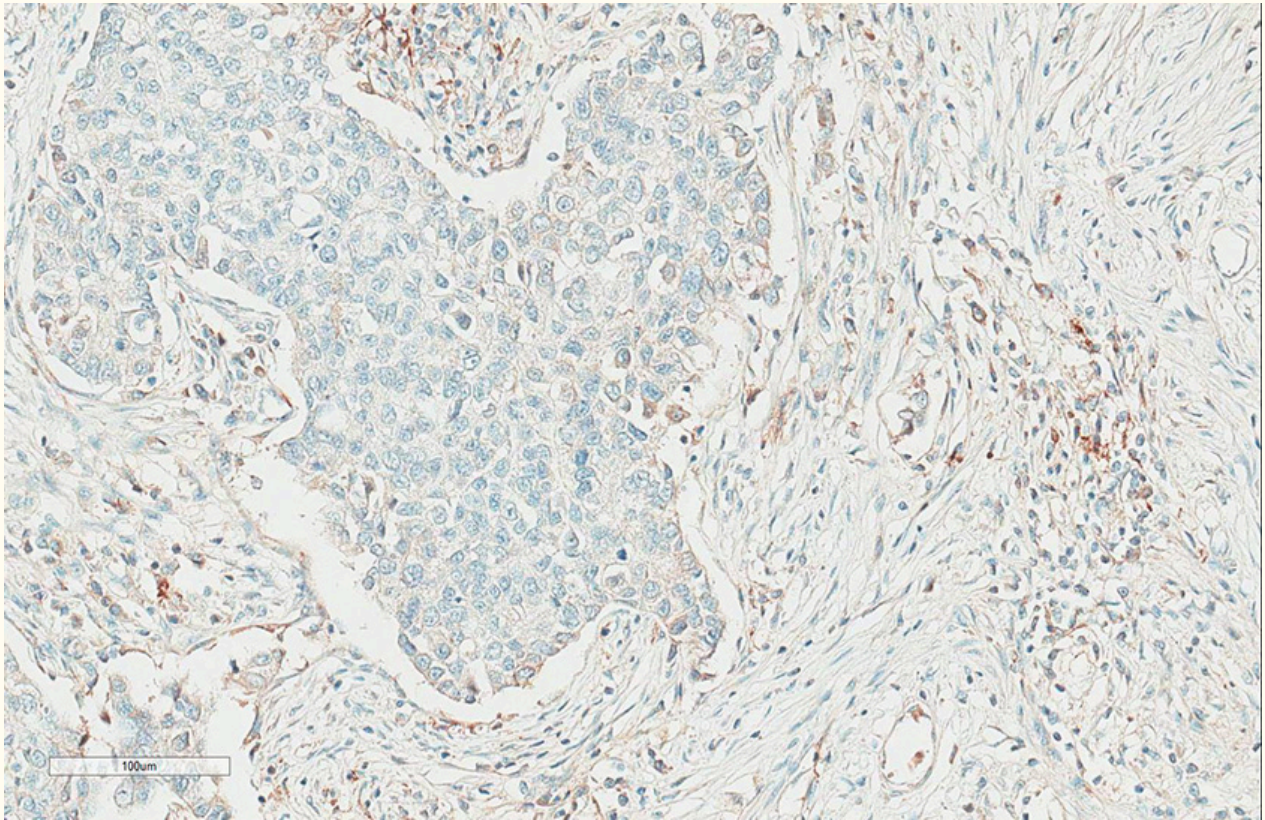
Figure 5B

Figure 5B caption

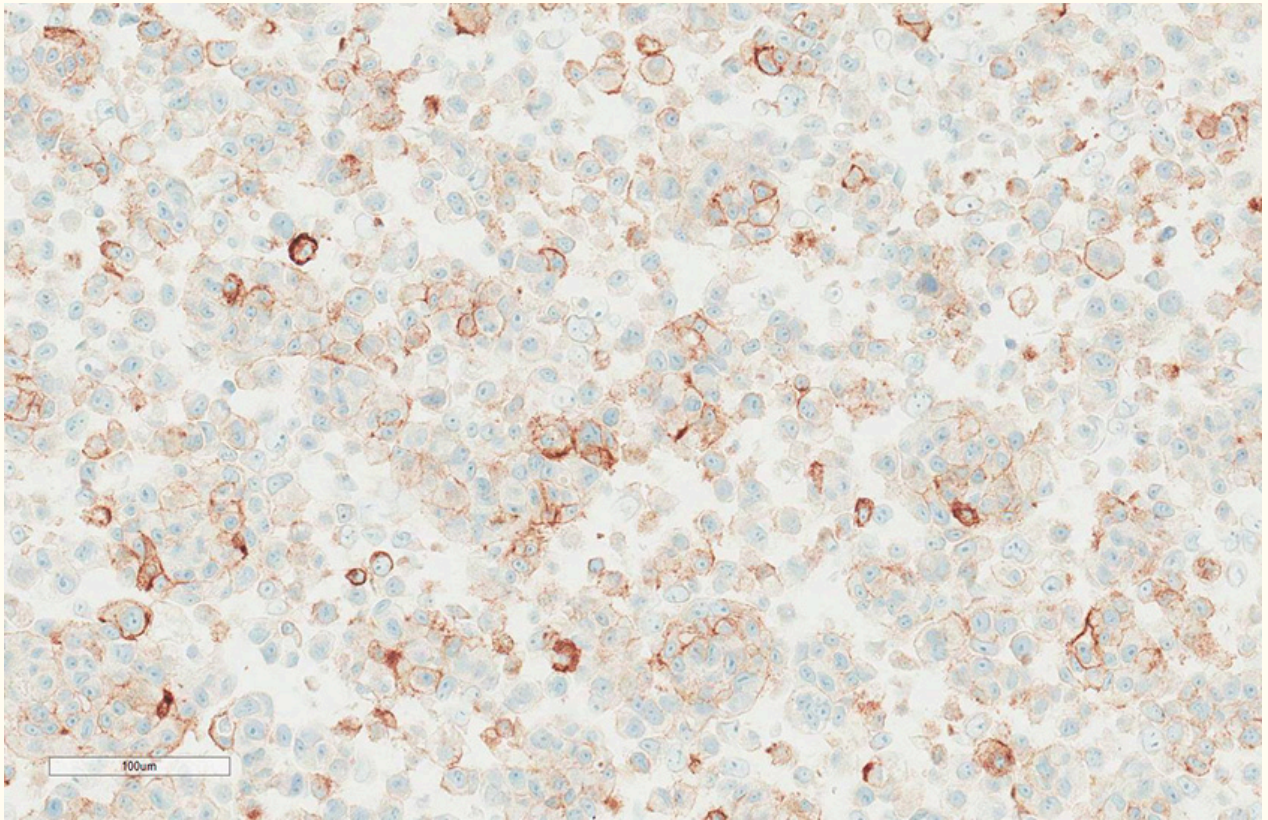Participant submission SP263 assay 'F'. Non-specific inappropriate staining (Mark 2/5)

Figure 6A
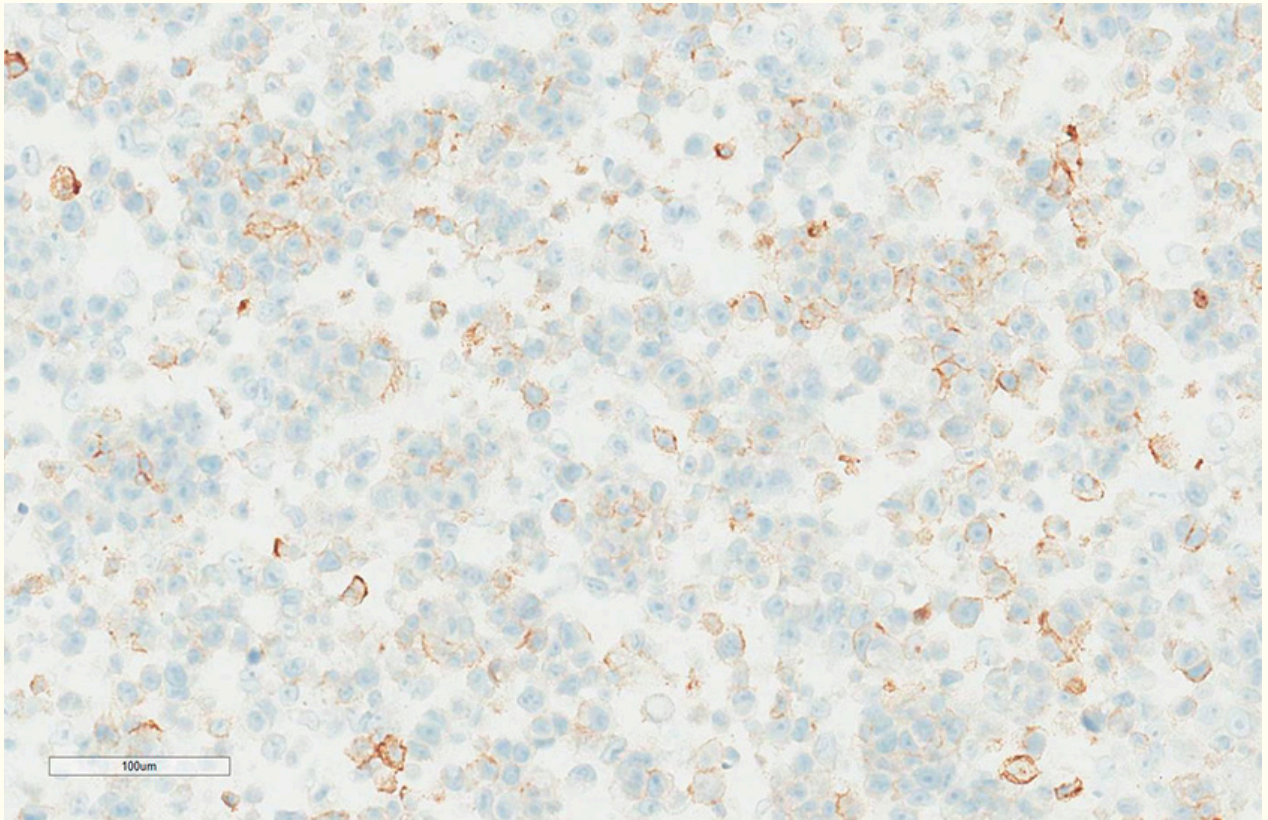
Figure 6A caption

Ventana/Roche SP263 Gold 'B' 50-79% TPS

Figure 6B

Figure 6B caption

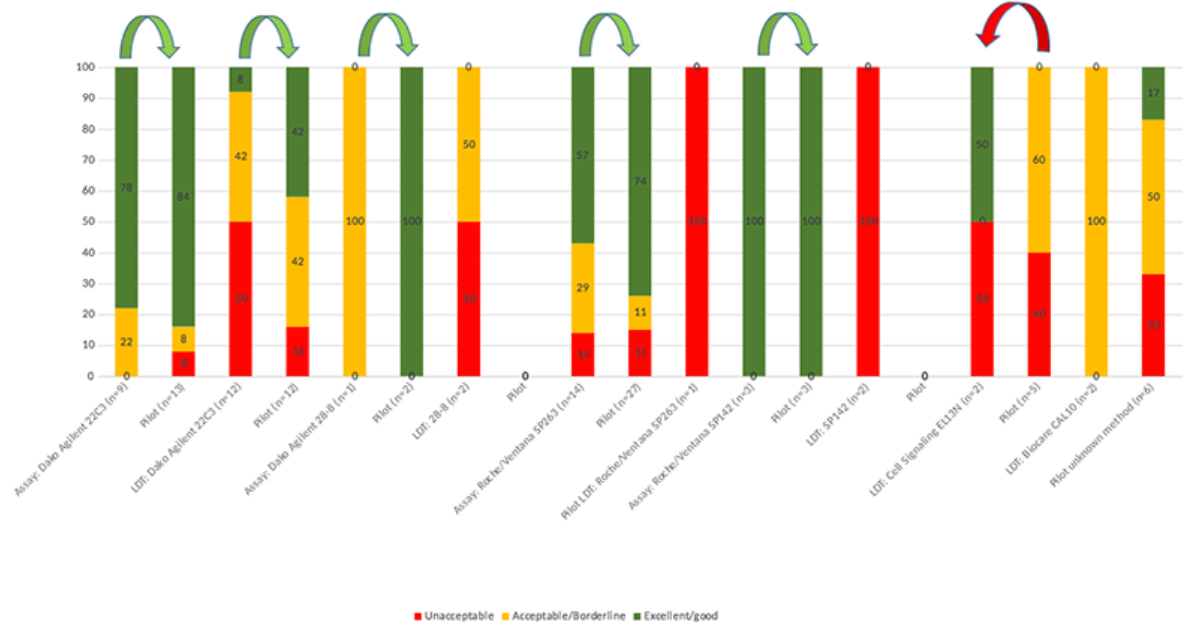Participant submission SP263 'B'. Interpreted as 25-49% by committee

Figure 7

Figure 7 caption

Comparison of pre-pilot results and pilot results.

Conclusions

There are multiple challenges in implementing a PD-L1 for NSCLC proficiency testing program, which are being experienced in EQA programs around the world [8,14,15,17]. The collaboration of RCPAQAP with UK NEQAS has been successful in providing meaningful results and recommendations to Australian laboratories in PD-L1 for non-small cell lung carcinoma, where participation rates were expected to be low. The pre-pilot PD-L1 meeting at UK NEQAS was successful in establishing assessment guidelines for PD-L1 assessment in NSCLC. Findings suggested that use of a clinically validated PD-L1 immunohistochemistry (IHC) assay performs better during assessment than adopting an LDT. However, devising and validating an optimal method against the clinical assay associated with the PD-1/PD-L1 therapy offered and continual verification of the test can produce the expected results [8]. Following the pre-pilot, it was recommended by UK NEQAS that an optimal in-house control for participants should include a dynamic range of PD-L1 expression on NSCLC in addition to a sample of tonsil [8]. Tonsil was preferred over placenta to portray varying levels of PD-L1 in normal tissue. RCPAQAP continues to collaborate with UK NEQAS and a technical EQA program for PD-L1 in NSCLC is now available for enrolment. The formation of IQNPath also proves the importance of collaboration between EQA professionals and industry to exchange expertise, ideas and promote interaction.

Original submitted files for images

Below are the links to the authors' original submitted files for images.

Original image file for Figure 1
Click here to view.

Original image file for Figure 2
Click here to view.

Original image file for Figure 3
Click here to view.

Original image file for Figure 5B
Click here to view.

Original image file for Figure 4A
Click here to view.

Original image file for Figure 6A
Click here to view.

Original image file for Figure 4B
Click here to view.

Original image file for Figure 6B
Click here to view.

Original image file for Figure 5A
Click here to view.

Original image file for Figure 7
Click here to view.

References

1. Casaluce F, Sgambato A, Sacco P, Palazzolo G, Maione P, Rossi A, et al. Emerging drugs targeting PD-1 and PD-L1: reality or hope?. Expert opin emerg drugs. 2014 Dec;19(4):557-569.

2. Kerr K, Nicolson M. Non-small Cell Lung cancer, PD-L1, and the pathologist. Arch Pathol Lab Med. 2016 Mar.140(3)

3. Zhang M, Li G, Wang Y, Wang Y, Zhao S, Haihong P, et al. PD-L1 expression in lung cancer and its correlation with driver mutations: a meta-analysis. Sci Rep. 2017 Aug 31;7(1):10255-.

4. Ettinger D, Wood D, Akerley W, Bazhenova LA, Borghaei H, Camidge DR, et al. NCCN Guidelines Insights: Non-Small Cell Lung Cancer, Version 4.2016. J Natl Compr Canc Netw. 2016 Mar;14(3):255-264.

5. Ameratunga M, Asadi K, Lin X, Walkiewicz M, Murone C, Knight S, et al. PD-L1 and Tumour Infiltrating Lymphocytes as Prognostic Markers in Resected NSCLC. PLoS One. 2016 Apr 22;11(4):e0153954-.

6. Alsaab H, Sau S, Alzhrani R, Tatiparti K, Bhise K, Kashaw SK, et al. PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome. Front Pharmacol. 2017 Aug 23;8:561-.

7. Topalian S, Drake C, Pardoll D. Targeting the PD-L1/B7-H1(PD-1) pathway to activate anti-tumor immunity. Curr Opin Immunol. 2012 Apr;24(2):207-12.

8. Newman A, Parry S, Wilkinson D, O'Grady T, Maxwell P, Allen D, et al. UK NEQAS ICC & ISH pre-pilot meeting for PD-L1.. IHC in NSCLC. .
Available from: http://www.ukneqasiccish.org/wp/wp-content/uploads/2017/09/PD-L1_prepilot_write_up_070917.pdf

9. Brahmer J, Reckamp K, Baas P, Crinò L, Eberhardt WE, Poddubskaya E, et al. Nivolumab versus docetaxel in advanced squamous non-small-cell lung cancer. N Engl J Med. 2015 Jul 9;373(2):123-135.

10. Garon E, Rizvi N, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. N Engl J Med. 2015 May 21;372(21):2018-2028.

11. Rebalatto M, Mistry A, Sabalos C, Walker J, Midha A, Steele K, et al. Development of a PD-L1 companion diagnostic assay for treatment with MED14736 in NSCLC and SCCHN patients [abstract 8033]. J Clin Oncol. 2015 May.33(suppl.)

12. Spira A, Park K, Mazieres J, Vansteenkiste J, Rittmeyer A, Ballinger M, et al. Efficacy, safety and predicitve biomarker results from a randomized phase II study comparing MPDL3280A vs docetaxel in 2L/3L NSCLC (POPLAR) [abstract 8010]. J Clin. Oncol. 2015 May.33suppl.)

13. Tsao B, Kerr K, Kockx M, Beasley MB, Borczuk AC, Botling J, et al. PD-L1 Immunohistochemistry Comparability Study in Real-Life Clinical Samples: Results of Blueprint Phase 2 Project. J Thorac Oncol. 2018 Sep;13(9):1302-1311.

14. NordiQC. Assessment Run C1 PD-L1. Aalborg, Denmark. 2017.
Available from: http://www.nordiqc.org/downloads/assessments/96_102.pdf

15. NordiQC. Assessment Run C2 PD-L1. Aalborg, Denmark. 2018.
Available from: http://www.nordiqc.org/downloads/assessments/100_102.pdf

16. A Quality Undertaking. Connecting EQA and industry to ensure quality diagnostic testing. The Pathologist. 2017 Oct.
Available from: https://thepathologist.com/issues/1017/a-quality-undertaking/

17. cIQc Non-small cell lung cancer EQA for PD-L1 educational run. .
Available from: http://nordiqc2017.dk/wp-content/uploads/3_PD-L1-for-Aalborg-2017-final.pdf

Copyright & License

VascularCell

PUBLIVERSE ONLINE

The present article has been published in Vascular Cell journal by Publiverse Online S.R.L.