The challenges of implementing a PD-L1 proficiency testing program

J. Pagliuso[1], S. Parry[2], Z. Haffajee[1], K Miller[2]

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

Background

One important initiative that commenced at the RCPAQAP in 2017 was collaboration with UK NEQAS ICC & ISH for the challenging implementation of a PD-L1 immunohistochemistry (IHC) proficiency testing program for non-small cell lung carcinoma (NSCLC). An RCPAQAP participant survey in 2016 showed that only eight laboratories were performing PD-L1 testing. **Challenge 1: Participation** It would not be viable or meaningful to establish an external quality assurance (EQA) program for eight participants. **Challenge 2: Multiple PD-L1 Biomarkers for multiple therapies** PD-L1 is unique to other biomarkers in that at least four different therapies have been developed or are in the development phase targeting the PD-1/PD-L1 pathway, and have been clinically validated with four different IHC antibody clones on different staining protocols **Challenge 3: Different scoring systems** These clones use different scoring systems and have different cut-off thresholds for defining positivity for the application of each drug. These were the same challenges encountered by EQA programs around the world.

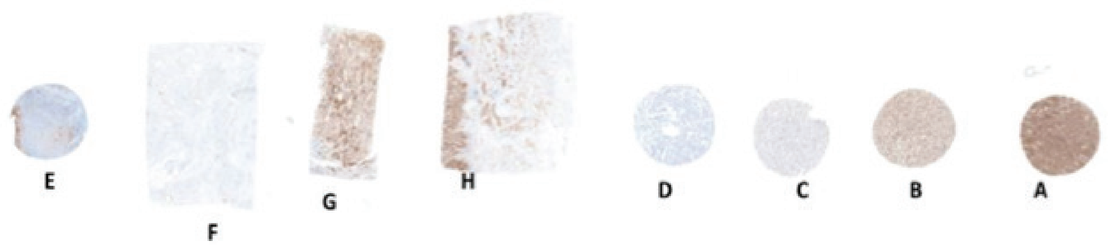
Aim

In 2016, UK NEQAS were also at the beginnings of establishing their own program. The aim of the collaboration with UK NEQAS was to increase the sample size of the pilot program to create meaningful results. UK NEQAS conducted a pre-pilot in early 2017 which attracted a total of 47 participants, including 13 Australian laboratories. The aim of the pre-pilot was to discuss the complexities surrounding PD-L1 submissions and establish a set of guidelines to help harmonise the assessment process.

Methods

Pre-pilot participants were sent 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.

Figure 1. Example of PD-L1 expression for multiblock 2



Participants were also asked to submit their methodology with the returned stained slides. The assessment was attended by 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. **Challenge 4: Varying clinical cut-offs for positivity.** A BIN system was established as shown in table 1.

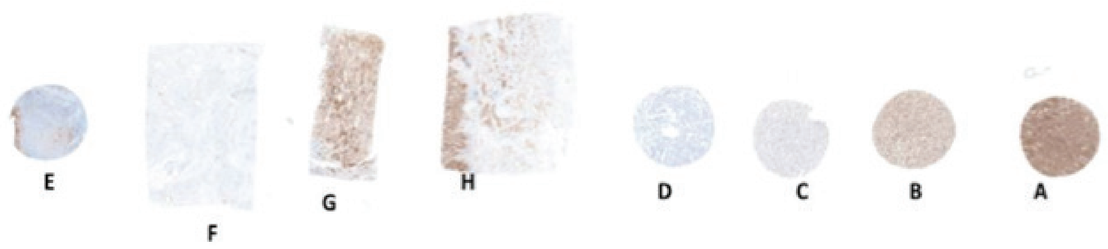
The use of the BIN system allowed a harmonised approach to the assessment of all scenarios of submitted slides. The BIN TPS was then applied to each core/section on the gold standard for each commercial assay as seen in figure 2.

Table 1. Tumour proportion score BIN categories

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 baseline to compare participant results.

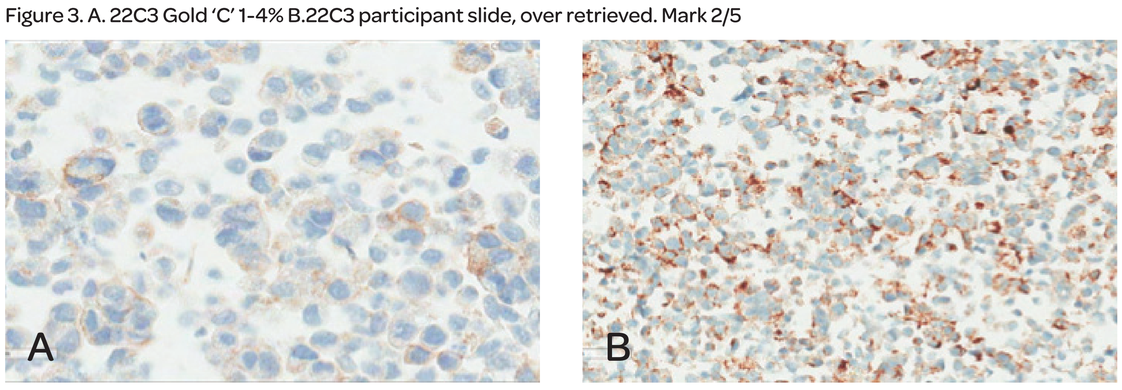
Figure 2. TPS BIN category applied to Gold sections



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.

Figure 3. A. 22C3 Gold 'C' 1-4% B.22C3 participant slide, over retrieved. Mark 2/5



The UK NEQAS distributed tonsil control tissue (sample E) was assessed as either acceptable, borderline or unacceptable.

Challenge 6: Laboratory Developed Tests (LDTs) 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. Twenty out of 47 participants submitted an LDT stained slide for the pre-pilot.

Results

Results of the pre-pilot are shown in table 2. The two most common approved assays used were Dako 22C3 (0/9 unacceptable) and Ventana SP263 (2/14 unacceptable). 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 Bond III) achieved a good/excellent result.

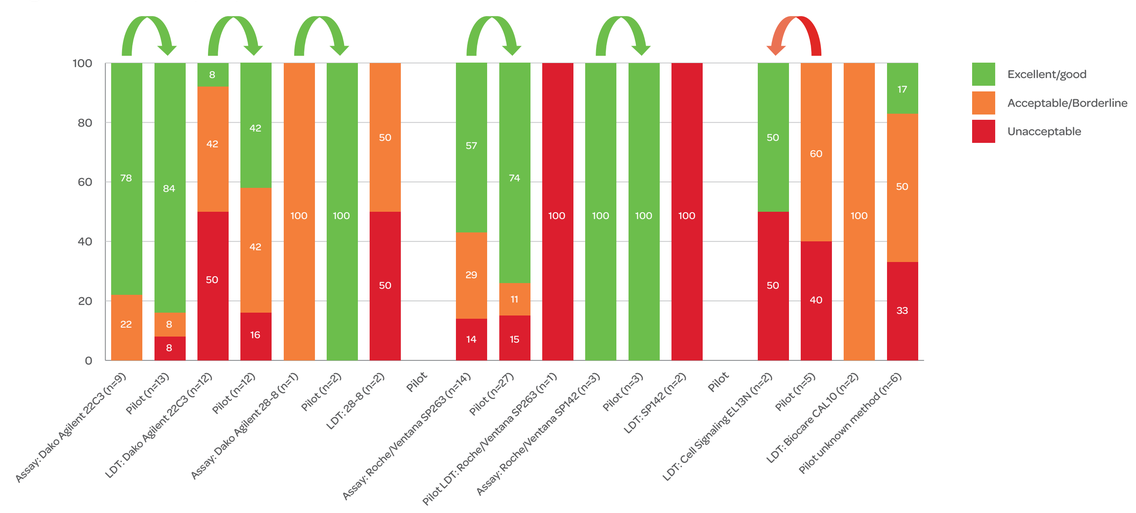
Table 2. Results of pre-pilot, n=47

PD-L1 Assay	Automation	Detection Kit	Good/ Excellent	Acceptable/ Borderline	Unacceptable	Total 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 BondMax	Leica Bond Polymer Refine	–	–	Dako	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 28-8 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 XT	Ventana Ultraview	–	–	1 (100%)	n=1
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

Three out of thirteen Australian participants received a good/excellent result. Seven out of thirteen results submitted from Australia were LDTs.

The pre-pilot was followed up with a pilot at the end of 2017. 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 **Challenge 7: Donation of tissue** plus two new laboratories who are regular donators to the RCPAQAP programs. General improvement was seen between the pre-pilot and pilot results, with the exception of the E1L3N clone as seen in figure 4.

Figure 4. Results of pre-pilot, n=47



Two out of the five Australian participants that received a score of two in the pre-pilot, received scores of four in the pilot. The remaining three Australian participants did not show any improvement and received a score of two in the pilot.

Challenge 8: Interpretation proficiency testing It should be noted that participants' interpretation of the TPS was not assessed as UK NEQAS ICC & ISH is purely a technical EQA scheme. Like many other EQA program providers, RCPAQAP is a member of IQNPath (International Quality Network for Pathology) who are creating an online digital self-assessment for pathologists to test TPS interpretation for the four FDA approved PD-L1 assays. RCPAQAP will keep participants updated with the progress of this implementation.

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[1,2,3]. The UK NEQAS pre-pilot PD-L1 meeting 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 that 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[4]. An optimal in-house control for participants would include a dynamic range of PD-L1 expression on NSCLC in addition to a sample of tonsil[4].

A technical EQA program for PD-L1 in NSCLC is now available for enrolment from the RCPAQAP website in collaboration with UK NEQAS ICC & ISH. The RCPAQAP will communicate progress on the IQNPath PD-L1 interpretation EQA program through myQAP.

References

1. NordiQC, Assessment Run C1 PD-L1 (2017) viewed at http://www.nordiqc.org/downloads/assessments/96_102.pdf
2. NordiQC, Assessment Run C2 PD-L1 (2018) viewed at http://www.nordiqc.org/downloads/assessments/100_102.pdf.
3. cIQc Non-small cell lung cancer EQA for PD-L1 educational run viewed at http://nordiqc2017.dk/wp-content/uploads/3_PD-L1-for-Aalborg-2017-final.pdf
4. UK NEQAS ICC & ISH pre-pilot meeting for PD-L1 iHC in NSCLC viewed at http://www.ukneqasiccish.org/wp/wp-content/uploads/2017/09/PD-L1_prepilot_write_up_070917.pdf.

RCPAQAP
The Royal College of Pathologists of Australasia
Quality Assurance Programs

FIND OUT MORE RCPAQAP.COM.AU
Email : anatomical.pathology@rcpaqap.com.au
NATA Accredited Proficiency Testing Scheme Provider. Number 14863 Accredited to ISO/IEC 17043:2010

UK NEQAS
Immunocytochemistry
& In-Situ Hybridisation