Strengths and Weaknesses of PD-L1 testing: Pathology perspective

Prof Keith M Kerr
Department of Pathology
Aberdeen University Medical School, Aberdeen Royal Infirmary
Foresterhill, Aberdeen
Disclosures

Consultancy

• AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Eli Lilly, Merck Serono, Merck Sharp & Dohme, Novartis, Pfizer, Roche, Ventana

Honoraria

• AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Eli Lilly, Merck Serono, Merck Sharp & Dohme, Novartis, Pfizer, Roche, Ventana
Agenda

➢ Background
➢ Strengths of PD-L1 testing
➢ Weaknesses of PD-L1 testing
➢ Conclusions
The Immune Response is a Very Complex Multifactorial ‘Reaction’

The PD-1/PD-L1 axis immune checkpoint in NSCLC

- A therapeutic target
- An important biomarker

NSCLC, non-small cell lung cancer; PD-1, programmed death 1; PD-L1, programmed death ligand 1.
Therapeutic Aims & Assumptions

• Inhibit the interaction of PD-1 and PD-L1
  • In order to ‘take the brakes off’ an existing tumour-specific immune response

• We hope this tumour-specific immune response actually exists
  – This requires immunogenicity
  – Which requires antigenicity (neoantigens) to be ‘visible’ to the immune system

• A surrogate for probable neoantigen load is tumour mutational burden (TMB)

Potential Biomarkers

Numbers of mutations in diagnostic panels

Whole tumour mutation burden: Whole Exome Sequencing

PD-1~PD-L1 Interaction is present

Specific immunity is available – the tumour is ‘inflamed’

THIS is the drug target!
The Cancer Immunogram

Numerous factors determine effectiveness of a tumour-directed immune response

Some cancers are NOT immunogenic; there is no immune response

How does the tumour EVADE the immune response?

Inhibitory immune checkpoints are one important mechanism
Why Do We Need Biomarkers for Immunotherapy?

Precision medicine is a reality for many tumour types

Avoidance of harm?
• There are toxicities from these drugs
• Is there a subgroup of patients who fair worse on I-O treatment?
• Alternative treatment would be better

Not all patients respond to and benefit from these treatments
• Enrich the treatment population for benefit
• *How many to treat, to get one response?*
• Benefit is relative to standard of care

Financial burden of expensive therapy
# Five most advanced PD1 axis inhibitors and their biomarker assays

<table>
<thead>
<tr>
<th>Drug</th>
<th>Company</th>
<th>PD-L1 Diagnostic Ab clone</th>
<th>Staining Platform</th>
<th>Clinically relevant cut offs *</th>
<th>Biomarker status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nivolumab</td>
<td>Bristol-Myers Squibb</td>
<td>28-8 (Dako)</td>
<td>Dako Link 48</td>
<td>TC ≥1%, 5%, 10%</td>
<td>Complementary</td>
</tr>
<tr>
<td>Pembrolizumab</td>
<td>Merck/MSD</td>
<td>22C3 (Dako)</td>
<td>Dako Link 48</td>
<td>TC ≥ 1, 50%</td>
<td>Companion</td>
</tr>
<tr>
<td>Atezolizumab</td>
<td>Genentech/Roche</td>
<td>SP142 (Ventana)</td>
<td>Ventana BenchMark ULTRA</td>
<td>TC ≥ 1, 10, 50% IC ≥ 1, 5, 10%</td>
<td>Complementary</td>
</tr>
<tr>
<td>Durvalumab</td>
<td>Astra-Zeneca</td>
<td>SP263 (Ventana)</td>
<td>Ventana Benchmark</td>
<td>TC ≥ 25%</td>
<td>Not sure</td>
</tr>
<tr>
<td>Avelumab</td>
<td>Pfizer/ Merck Serono</td>
<td>73-10 (Dako)</td>
<td>Dako Link 48</td>
<td>TC ≥ 1, 50, 80%</td>
<td>Unknown</td>
</tr>
</tbody>
</table>
Strengths of PD-L1 testing

• It works as a selective biomarker to some extent
• It is easy to access and carry out the test relatively
• It is quick to read and generate the ‘result’ relatively
• It is cheap
• Test outcomes have potential to be reliable
• It is the only biomarker with ‘approvals’ for clinical usage
The test ‘works’: in Advanced NSCLC, PD-L1 Expression by Immunohistochemistry Can Enrich for Response

- Response rates in the ITT population are 14% to 20%
- Response rates in ‘selected’ populations are 30% to 45% (can be higher)
- RR translates into survival benefit
- Most hazard ratios for benefit from PD-1 axis IO improve as PD-L1 expression levels increase
- The required ‘power’ of the biomarker to improve outcome depends upon the activity of the current standard of care comparator
The test: Ease of access and conduct

• Immunohistochemistry is commonplace in pathology labs
  • Familiarity
  • Skill sets available
  • Routine procedure

• Ease of Access..........if
  • An appropriate staining platform is available
  • Funding is available for Trial validated assays on the market
Quick to read and generate the ‘result’

- IHC is generally a ‘next day’ test assuming it is carried out in house
- The individual case is usually relatively quick to read
  - Workload allowance circa 20 mins
- Potential for rapid communication thereafter
Cheap!

• IHC reagents for the trial validated assay ➢ 70 – 100 Euro
• Scientific staff and Pathologist time ➢ Modest!
• Modelled cost to NHS in Scotland ➢ Approx 180 Euro

• Possible biomarker competitors ➢ Thousands
BluePrint2: Strong reliability among all pathologists on tumor cell scoring

<table>
<thead>
<tr>
<th></th>
<th>DIGITAL</th>
<th>GLASS SLIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All cases</td>
<td>NSCLC tissue only</td>
</tr>
<tr>
<td>22C3</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>28-8</td>
<td>0.86</td>
<td>0.88</td>
</tr>
<tr>
<td>SP-142</td>
<td>0.81</td>
<td>0.85</td>
</tr>
<tr>
<td>SP-263</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>73-10</td>
<td>0.89</td>
<td>0.91</td>
</tr>
<tr>
<td>All assays</td>
<td>0.91</td>
<td>0.93</td>
</tr>
</tbody>
</table>

|                | All cases        | NSCLC tissue only | Cytology only |
| 22C3           | 0.89             | 0.87              | 0.88          |
| 28-8           | 0.92             | 0.94              | 0.87          |
| SP-142         | 0.86             | 0.84              | 0.90          |
| SP-263         | 0.86             | 0.89              | 0.79          |
| 73-10          | 0.93             | 0.93              | 0.84          |
| All assays     | 0.86             | 0.89              | 0.77          |

Fleiss Kappa: >0.90: excellent 0.75-0.9: good

Tsao MS, Kerr KM, Hirsch FR et al, Blueprint 2 WCLC 2017

Weaknesses of PD-L1 testing

• It works as a selective biomarker............to some extent
  • But it isn’t perfect!
• It is easy to access and carry out the test...........relatively
  • Dependency on laboratory availability of platform
• It is quick to read and generate the ‘result’...........relatively
• It is cheap........
  • But everything is relative
• Test outcomes have potential to be reliable
  • However...........
• It is the only biomarker with ‘approvals’ for clinical usage but it is IHC
  • and there is lots of work ongoing to challenge this position
For NSCLC, PD-L1 IHC Has the Ability to Enrich for Response and Treatment Benefit, however......

PD-L1 IHC is not a perfect biomarker

• Not all ‘positive’ patients respond
• Occasional responders in ‘biomarker negative’ populations
• But no biomarker is perfect
• Expectations in the real, clinical world
Why is PD-L1 IHC imperfect?

Possible Confounders of *Tumour Cell* PD-L1 IHC Prediction

- Heterogeneity of expression—potential for sampling error
- Cutoffs in a biological continuum
Cross-validation of PD-L1 scoring: TMAs vs Whole Sections

Table 1: Agreement in the % of PD-L1 positive neoplastic cells using TMAs vs Whole Sections (No of cases, Total N=237)

<table>
<thead>
<tr>
<th>PD-L1 in TMA's</th>
<th>&lt;1%</th>
<th>1-&lt;5%</th>
<th>5-&lt;10%</th>
<th>10-&lt;25%</th>
<th>25-&lt;50%</th>
<th>≥50%</th>
<th>Not Evaluable</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1%</td>
<td>89</td>
<td>17</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1-&lt;5%</td>
<td>5</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5-&lt;10%</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10-&lt;25%</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>25-&lt;50%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>≥50%</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>Not Evaluable</td>
<td>16</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

Small samples tend to underestimate PD-L1 expression

Setting ‘Whole Section’ result as ‘Gold Standard’:

<table>
<thead>
<tr>
<th>Cut-off for PD-L1 positivity</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>78.5%</td>
<td>89.9%</td>
</tr>
<tr>
<td>5%</td>
<td>84.7%</td>
<td>92.5%</td>
</tr>
<tr>
<td>50%</td>
<td>79.1%</td>
<td>98.2%</td>
</tr>
</tbody>
</table>
Binary Output Versus Biological Continuum

Biomarker is ABSENT
You are unlikely to benefit from therapy

THRESHOLD CUTOFF

Biomarker is ABSENT or at a LOW LEVEL
You are unlikely to benefit from therapy

Biomarker is PRESENT at an INTERMEDIATE LEVEL
You may benefit from therapy

Biomarker is PRESENT at a HIGH level
You are likely to benefit from therapy

Biomarker is PRESENT
You are likely to benefit from therapy

1%

50%

80%

Biological continuum of biomarker expression
Relationship between PD-L1 expression and outcome

Nivolumab Plus Ipilimumab in First-line NSCLC: Efficacy Across All Tumor PD-L1 Expression Levels

Presented by Matthew Hellmann at 2016 ASCO Annual Meeting

Prevalence of PD-L1 Positivity and ORR by Quartiles of PD-L1 Proportion Score

Prevalence, all screened patients: 332 (39.2%), 255 (31.0%), 55 (6.7%), 71 (8.6%), 120 (14.6%)

ORR in CTA-evaluable patients: 7 (8.1%), 19 (12.9%), 6 (19.4%), 13 (29.6%), 39 (45.4%)

*Prevalence and ORR (RECIST v.1.1 by central review) assessed in patients whose samples were evaluable by the CTA, regardless of the interval between cutting and staining.

Analysis cut-off date: August 29, 2014.
Other possible Confounders of *Tumour Cell PD-L1 IHC Prediction*

- Expression on tumour cells (TC) and/or immune cells (IC)
  - What about immune cell scoring?
- Intrinsic induction of PD-L1: oncogenic pathway driven
- Other immune regulatory mechanisms are also active
  - PD-L1 mediated regulation is not ‘unique’
- Technical issues
  - Does the assay correctly identify patients?
  - Did the assay work this time?
BluePrint2: Poor reliability among **all pathologists** on **immune cell** scoring

<table>
<thead>
<tr>
<th></th>
<th>DIGITAL</th>
<th>GLASS SLIDES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All cases</td>
<td>NSCLC tissue only</td>
</tr>
<tr>
<td>22C3</td>
<td>0.28</td>
<td>0.23</td>
</tr>
<tr>
<td>28-8</td>
<td>0.19</td>
<td>0.14</td>
</tr>
<tr>
<td>SP-142</td>
<td>0.36</td>
<td>0.28</td>
</tr>
<tr>
<td>SP-263</td>
<td>0.25</td>
<td>0.13</td>
</tr>
<tr>
<td>73-10</td>
<td>0.17</td>
<td>0.10</td>
</tr>
<tr>
<td>All assays</td>
<td>0.19</td>
<td>0.11</td>
</tr>
</tbody>
</table>

|               | All cases        | NSCLC tissue only |
| 22C3          | 0.27             | 0.19              |
| 28-8          | 0.29             | 0.19              |
| SP-142        | 0.33             | 0.25              |
| SP-263        | 0.17             | 0.10              |
| 73-10         | 0.17             | 0.11              |
| All assays    | 0.21             | 0.13              |

Fleiss Kappa: 0.40-0.59: weak
0.20-0.39: minimal
<0.01-0.20: slight/none

Tsao MS, Kerr KM, Hirsch FR et al, Blueprint 2 WCLC 2017

Intrinsic induction of PD-L1 expression

- Upregulation related to oncogenic pathway activation in the tumour
  - JAK-STAT
  - PI3K
  - Probably others due to crosstalk

Implies PD-L1 expression that may not be immunologically active

We have little idea, so far, how to identify these cases

Teng MWL et al. Cancer Res 2013
Other possible factors impacting therapy response

Host factors
- General immune status
- Gut microbiome

Tumour Micro-environment
- Other checkpoints
- Soluble immune inhibitors
- Inhibitory cell populations
- Inhibitory tumour metabolism
- Tumour ‘invisible’ or impervious to immune killing

PD1~PD-L1 inhibition is active

There is an anti-tumour immune response to re-activate

The tumour is immunogenic

‘The Lottery Model’
Blank CU et al. Science 2016
Ease of access to platform and trial validated assay

• Your lab does not have the correct platform
• Your lab does not have the budget
Five drug- PD-L1 IHC assay combinations?

• Trial validated assay for each drug?
  • It is unlikely that labs will provide multiple tests
  • Staining platform availability?
• One Trail validated assay for all drugs?
• Lab Developed Test?
• Laboratory developed tests (LDTs)
  • ‘Modification’ of the companion/complementary assay
  • LDT built around a ‘trial’ clone
  • LDT built around another clone

<table>
<thead>
<tr>
<th>Drug</th>
<th>Assay based on clone</th>
<th>Commercial platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nivolumab</td>
<td>28-8</td>
<td>Dako</td>
</tr>
<tr>
<td>Pembrolizumab</td>
<td>22C3</td>
<td>Dako</td>
</tr>
<tr>
<td>Atezolizumab</td>
<td>SP142</td>
<td>Ventana</td>
</tr>
<tr>
<td>Durvalumab</td>
<td>SP263</td>
<td>Ventana</td>
</tr>
<tr>
<td>Avelumab</td>
<td>73-10</td>
<td>Dako</td>
</tr>
</tbody>
</table>

Other platforms are available!

Outcome of an IHC test is a function of the primary antibody AND the detection system used
Blueprint, NCCN and French Studies: Tumour cell staining

- **Blueprint 1**: 28-8, 22C3 and SP263 similar across range.
- **NCCN study**: SP142 stained fewer tumour cells; 73-10 stained more tumour cells.
- **French study**: Similar data from ‘AZ 500’ study

Ratcliffe M et al CCR 2017

LDTs? Issues!

Hirsch et al. JTO 2016; Rimm DL et al. JAMA Oncol 2017; Adam J et al (in press); Tsao MS et al WCLC 2017
The Trial validated assays are (relatively) expensive

• Reimbursement of IHC testing in some countries is incompatible with the costs of Companion or Complementary Commercial Kit assays
Reliability of the test outcome?

• The test did not this work this time around
  • Pre-analytics
  • Tissue sample is inadequate or exhausted
  • Something ‘happened’ !! (it does............)

• The test never works properly, but you don’t realize............

• The pathologist is at fault
  • Training
  • Bad day etc............
Selection of LDT

<table>
<thead>
<tr>
<th></th>
<th>Dako</th>
<th>Ventana</th>
<th>Leica</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center 1</td>
<td>Ref.</td>
<td>0.94</td>
<td>0.79</td>
<td>0.8</td>
</tr>
<tr>
<td>Center 2</td>
<td>28-8</td>
<td></td>
<td></td>
<td>28-8 (center 1)</td>
</tr>
<tr>
<td>Center 3</td>
<td>22C3</td>
<td></td>
<td></td>
<td>22C3 (center 1)</td>
</tr>
<tr>
<td>Center 4</td>
<td>SP263</td>
<td>0.83</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>Center 5</td>
<td>0.91</td>
<td>0.82</td>
<td></td>
<td>SP263 (center 5)</td>
</tr>
<tr>
<td>Center 6</td>
<td>0.81</td>
<td>0.77</td>
<td></td>
<td>SP263 (center 5)</td>
</tr>
<tr>
<td>Center 7</td>
<td>0.58</td>
<td>0.61</td>
<td>0.43</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.38</td>
<td>0.45</td>
<td>0.75</td>
</tr>
<tr>
<td>Selected LDT:</td>
<td>Dako: E1L3N, SP263</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ventana: 28-8, 22C3, E1L3N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leica: E1L3N, SP142, SP263</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard set for ‘comparability’ for Tumour Cell staining for individual tests was kappa >0.75

13/27 (48%) of LDTs failed to make the grade
PD-L1 IHC EQA data show there ARE issues with LDTs

**NORDIC QC**

- Labs using trial validated assay
  - 22C3, 28-8, SP263-based assay
  - 77-92% pass rate

- Labs using LDTs
  - 22C3 or E1L3N clones used
  - 20% pass rate

**UK NEQAS**

<table>
<thead>
<tr>
<th>Test Type</th>
<th>Median Score</th>
<th>Pass Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial validated Assays</td>
<td>4.5</td>
<td>100%</td>
</tr>
<tr>
<td>Laboratory developed tests (LDTs)</td>
<td>2.5</td>
<td>35% Borderline; 65% Failed</td>
</tr>
</tbody>
</table>

Data courtesy of Dr Keith Miller

It is the only biomarker with ‘approvals’ for clinical usage but it is IHC..................

and there is lots of work ongoing to challenge this position

• IHC has a questionable reputation in the oncology community?
• It is a subjective assessment
• Other sexier, ‘more scientific’, expensive molecular tests are coming along..........
T-effector Signature Does Not Perform Better than PD-L1 IHC

**T\text{eff}**

<table>
<thead>
<tr>
<th></th>
<th>ITT-WT</th>
<th>Teff high (WT)</th>
<th>Teff low (WT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>692 (87%)</td>
<td>284 (43%)</td>
<td>374 (57%)</td>
</tr>
</tbody>
</table>

**PD-L1 IHC**

<table>
<thead>
<tr>
<th></th>
<th>TC2/3 or IC2/3 (WT)</th>
<th>TC1/2/3 or IC1/2/3 (WT)</th>
<th>TC0 and IC0 (WT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>244 (35%)</td>
<td>354 (51%)</td>
<td>338 (49%)</td>
</tr>
</tbody>
</table>

*Signature predictive ability is similar to PD-L1 expression*

---

*Stratified HRs for ITT, ITT-WT and Teff-high WT populations; unstratified HRs for all other subgroups.*

Conclusions

PD-L1 immunohistochemistry
• is a validated and approved biomarker
• It is the only ‘drug specific’ biomarker
• Clinically it can be reliably delivered
• It does enrich for treatment benefit

• But it is not perfect
• Implementation can be complicated
• Skiing off piste is attractive but potentially dangerous

• It is here to stay but may be combined with other factors
Thank you for listening!