Quality control of PD-L1 in NSCLC
- Results of two German ring trials
Step-wise approach

Predictive value

- Biological significance
- Technical aspects
  - Material
  - Assays
  - Scoring
Step-wise approach

1.) Internal ring trial (DGP\(^1\)) -

- Gain experience; identify challenges
  - 1a, Scoring: Observer-concordance
  - 1b, Staining: Laboratory-concordance
  - Assays: Comparisons

2.) Open ring trial (QuIP GmbH\(^2\)) -

- Quality assessment
  - Workshop
  - Open ring trial

---

1 'DGP' = German Association for Pathology
https://www.pathologie-dgp.de/

2 'QuIP' = 'Quality-Initiative Pathology'
= German EQA Institution
http://www.quip-ringversuche.de
1a) Internal ring trial: Scoring

1a: Internal ring trial,
PD-L1 Scoring
## PD-L1 Scoring

### Table A

<table>
<thead>
<tr>
<th>Assay, Antibody</th>
<th>Cell Type</th>
<th>Negative</th>
<th>Low/Weak</th>
<th>Medium</th>
<th>High/Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nivolumab (α-PD1; BMS)</td>
<td>Tumor</td>
<td>0-1%</td>
<td>1-5%</td>
<td>5-10%</td>
<td>≥10%</td>
</tr>
<tr>
<td>Pembrolizumab (α-PD1; MSD)</td>
<td>Tumor</td>
<td>0-1%</td>
<td>1-50%</td>
<td></td>
<td>≥50%</td>
</tr>
<tr>
<td>Atezolizumab (α-PD-L1; Roche)</td>
<td>Tumor</td>
<td>0-1%</td>
<td>1-5%</td>
<td>5 - 50%</td>
<td>≥50%</td>
</tr>
<tr>
<td>Atezolizumab (α-PD-L1; Roche)</td>
<td>Immune</td>
<td>0-1%</td>
<td>1-5%</td>
<td>5-10%</td>
<td>≥10%</td>
</tr>
<tr>
<td>Durvalumab (α-PD-L1; AstraZeneca)</td>
<td>Tumor</td>
<td></td>
<td>1-25%</td>
<td></td>
<td>≥25%</td>
</tr>
<tr>
<td>Avelumab (α-PD-L1; Pfizer + Merck)</td>
<td>Tumor</td>
<td>0-1%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table B

<table>
<thead>
<tr>
<th>Proportion-Score</th>
<th>Negative</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category:</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Cut-Off:</td>
<td>&lt;1%</td>
<td>≥1%</td>
</tr>
<tr>
<td>Intervall:</td>
<td>0 - 1%</td>
<td>≥1%</td>
</tr>
</tbody>
</table>

- Various cut-off criteria (assays; line of treatment)
- Criteria can be integrated into 6-step score
- Currently most important:
  - ≥50% and ≥1% stained tumor cells (TC)

1a) Internal ring trial: Scoring

1\textsuperscript{st} round, 'Training Set':
- 15 cases NSCLC
- Central staining
- 2 LDTs
- 9 Observer

2\textsuperscript{nd} round, 'Validation Set':
- 15 cases NSCLC
- Central staining
- 4 Clinical trial assays
- 9 Observer
1a) Interobserver-Concordance

<table>
<thead>
<tr>
<th>PD-L1 IHC</th>
<th>Light's kappa [95% CI], Tumor cell proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6-step score</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Training Set</strong></td>
<td></td>
</tr>
<tr>
<td>Lab dev. Assays</td>
<td></td>
</tr>
<tr>
<td>n=15 cases NSCLC</td>
<td></td>
</tr>
<tr>
<td>E1L3N on Leica</td>
<td>0.50 [0.37 - 0.64]</td>
</tr>
<tr>
<td>SP142 on Leica</td>
<td>0.49 [0.34 - 0.66]</td>
</tr>
<tr>
<td><strong>Validation Set</strong></td>
<td></td>
</tr>
<tr>
<td>Clinical Trial Assays</td>
<td></td>
</tr>
<tr>
<td>n=15 cases NSCLC</td>
<td></td>
</tr>
<tr>
<td>Dako 28-8</td>
<td>0.49 [0.36 - 0.63]</td>
</tr>
<tr>
<td>Dako 22C3</td>
<td>0.47 [0.34 - 0.63]</td>
</tr>
<tr>
<td>Ventana SP142</td>
<td>0.47 [0.35 - 0.62]</td>
</tr>
<tr>
<td>Ventana SP263</td>
<td>0.47 [0.34 - 0.63]</td>
</tr>
</tbody>
</table>

Interobserver-Concordance (TCs)

- 6-step score: $\kappa \approx 0.5$
- Cut-offs: $\kappa \approx 0.6 - 0.8$
- Interobserver-Concordance comparable between assays, LDTs

Scheel AH et al. Modern Pathology 2016. 29(10):1165-72
1b: Internal ring trial, PD-L1 Staining
1b) Internal ring trial: Staining

- NSCLC-TMA, 2mm cores, n=21
- Sections cut at central lab
- Staining and scoring at local labs
  - 10 institutes
  - 4 clinical trial assay (CTAs)
    (22C3, 28-8, SP263, SP142)
  - 11 lab-developed tests (LDTs)
  - Total: 27 stained sections
- Central review of stainings
- Controls: Cell-lines; Tonsil
1b) Interlaboratory-Concordance

<table>
<thead>
<tr>
<th>IHC: PD-L1</th>
<th>Sites (n)</th>
<th>Readable TMAcores</th>
<th>Light's kappa (±SD), Tumor proportion score</th>
<th>Proportion cut-off</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>6-step score</td>
<td>3-step score</td>
</tr>
<tr>
<td>22C3, Kit</td>
<td>3</td>
<td>90% (57/63)</td>
<td>0.69 (±0.09)</td>
<td>0.83 (±0.09)</td>
</tr>
<tr>
<td>28-8, Kit</td>
<td>3</td>
<td>94% (59/63)</td>
<td>0.66 (±0.17)</td>
<td>0.80 (±0.15)</td>
</tr>
<tr>
<td>SP263, Kit</td>
<td>4</td>
<td>81% (68/84)</td>
<td>0.66 (±0.22)</td>
<td>0.89 (±0.18)</td>
</tr>
<tr>
<td>SP142, Kit</td>
<td>6</td>
<td>90% (114/126)</td>
<td>0.63 (±0.16)</td>
<td>0.73 (±0.11)</td>
</tr>
<tr>
<td>LDAs</td>
<td>11</td>
<td>82% (189/231)</td>
<td>0.43 (±0.15)</td>
<td>0.50 (±0.18)</td>
</tr>
</tbody>
</table>

*High SD because most samples were ≥1%

Interlab-Concordance (TCs)
- 6-step score: \( \kappa \approx 0.66 \)
- Cut-offs: \( \kappa \approx 0.71 - 1.0 \)
- Interobserver-Concordance comparable between assays
- LDTs slightly lower

Interpretation of kappa's coefficient:
- \( \kappa < 0 \): poor
- \( \kappa = 0 - 0.2 \): slight
- \( \kappa = 0.21 - 0.4 \): fair
- \( \kappa = 0.41 - 0.6 \): moderate
- \( \kappa = 0.61 - 0.8 \): substantial
- \( \kappa = 0.81 - 1.0 \): (almost) perfect

Landis JT, Koch GG.
Biometrics 1977. 33(1);159-174

Scheel AH et al. submitted
1b) Internal ring trial: Interassay

1b, Interassay-Comparison
### NSCLC-TMA, CTAs

<table>
<thead>
<tr>
<th>Clinical trial assay:</th>
<th>22C3</th>
<th>28-8</th>
<th>SP263</th>
<th>SP142</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
<td><img src="image4" alt="Image" /></td>
</tr>
<tr>
<td>B1</td>
<td><img src="image5" alt="Image" /></td>
<td><img src="image6" alt="Image" /></td>
<td><img src="image7" alt="Image" /></td>
<td><img src="image8" alt="Image" /></td>
</tr>
<tr>
<td>D5</td>
<td><img src="image9" alt="Image" /></td>
<td><img src="image10" alt="Image" /></td>
<td><img src="image11" alt="Image" /></td>
<td><img src="image12" alt="Image" /></td>
</tr>
</tbody>
</table>
LDTs: QC+ and QC-

$n=11$ TMA-sections stained with LDTs

Employed Abs:
- 22C3 $n=4$
- 28-8 $n=3$
- SP263 $n=2$
- E1L3N $n=1$
- QR1 $n=1$

Manual review

Staining similar to 22C3 / 28-8 pharmDx = QC+, $n=6$

Employed Abs:
- 22C3 $n=1$
- 28-8 $n=1$
- SP263 $n=2$
- E1L3N $n=1$
- QR1 $n=1$

Staining divergent = QC-, $n=5$

Employed Abs:
- 22C3 $n=3$
- 28-8 $n=2$

Image-analysis / quantitation
NSCLC-TMA, LDTs

<table>
<thead>
<tr>
<th>Lab-dev. test</th>
<th>Core B2</th>
<th>Site 03</th>
<th>Site 04</th>
<th>Site 06</th>
<th>Site 04</th>
</tr>
</thead>
<tbody>
<tr>
<td>28-8 (QC+)</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td>E1L3N (QC+)</td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
</tr>
<tr>
<td>28-8 (QC-)</td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td>22C3 (QC-)</td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Tonsil
Quantitation by image-analysis

IHC: PD-L1

Hematoxylin

DAB

Positive mask

- Per core: analysis of 0.46mm² @ 200x
- Invasive carcinoma manually selected
- Color-deconvolution
- Read-outs: Hem. intensity, DAB-area relative to area of invasive carcinoma

Manual selection

Color-deconvolution
Image-analysis

- Bars: Mean of sites; antennae: SD
- QC+: Similar to 28-8 / 22C3; QC-: Less DAB

Scheel et al, Submitted
Summary internal ring trial

1. Reproducible PD-L1 scoring is achievable
2. Reproducible PD-L1 staining is achievable
   - Assays ('Kits') quite reliable
   - LDTs possible, calibration critical
   - Different primary antibody possible:
     28-8, 22C3, SP263, E1L3N, others

Results match other studies, in particular
Adams J., / InCA-Network; WCLC 2016, Vienna
2.) Open ring trial / QuIP#

2: Open ring trial / QuIP#-EQA

# 'QuIP' = 'Quality-Initiative Pathology'
= German EQA Institution
http://www.quip-ringversuche.de
Workshop PD-L1 in NSCLC,
Charité, Berlin 09/2016

- Lectures & live microscopy
- PD-L1 scoring criteria
- Challenges in scoring:
  - Cell type
  - Inclusion of low-intensity
  - Heterogeneity
- Assay-calibration (LDTs)
Open ring trial; work-flow

Case selection
- 3 lead institutes
- 14 cases
- Resection specimens

Case verification
- 5 panel institutes
- IHC: CTAs and LDTs
- Concordance ≥ 7/8

Ring trial
- 10 NSCLC cases
- 2 unstained slides / case
- Participants: n=87
Open ring trial; evaluation

Per case
- Correct score          2 points
- Not evaluable         1 point

Total
- Optimal               20 points
- 'Successful'          ≥ 18 points

Optional
- Return of stained slides for review
Open ring trial; cases, results

Composition

- NSCLC resection specimens, 9 cases
  (One case excluded; 'equivocal')
- Category 0, TC < 1%: 2 cases
- Category 1, TC 1%-49%: 2 cases
- Category 2, TC 50-100 %: 5 cases
- Results-form: Category, not percentage

Results

- Returned scores: 83
- participated successful: 60 (72%)
- participated: 23 (28%)

QuIP GmBH, used with permission
<table>
<thead>
<tr>
<th>Clone</th>
<th>Participants (%)</th>
<th>participated successful</th>
<th>participated</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1L3N</td>
<td>25 (30%)</td>
<td>22 (88%)</td>
<td>3 (12%)</td>
</tr>
<tr>
<td>28-8</td>
<td>20 (24%)</td>
<td>13 (65%)</td>
<td>7 (35%)</td>
</tr>
<tr>
<td>22C3</td>
<td>13 (16%)</td>
<td>9 (69%)</td>
<td>4 (31%)</td>
</tr>
<tr>
<td>QR1</td>
<td>6 (7%)</td>
<td>4 (67%)</td>
<td>2 (33%)</td>
</tr>
<tr>
<td>SP263</td>
<td>5 (6%)</td>
<td>2 (40%)</td>
<td>3 (60%)</td>
</tr>
<tr>
<td>Cal10</td>
<td>4 (5%)</td>
<td>3 (75%)</td>
<td>1 (25%)</td>
</tr>
<tr>
<td>ZR3</td>
<td>3 (4%)</td>
<td>2 (67%)</td>
<td>1 (33%)</td>
</tr>
<tr>
<td>SP142</td>
<td>2 (2%)</td>
<td>2 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>n.a.</td>
<td>5 (6%)</td>
<td>2 (40%)</td>
<td>3 (60%)</td>
</tr>
</tbody>
</table>
# Deviating scores, stainings

## A

<table>
<thead>
<tr>
<th>Clone</th>
<th>Lower</th>
<th>Higher</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1L3N</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>28-8</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>22C3</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>QR1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP263</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Cal10</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>ZR3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SP142</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td><strong>19</strong></td>
<td><strong>22</strong></td>
</tr>
</tbody>
</table>

## B

- Slide-sets returned: n = 18
- Correct staining / score: n = 1
- Scoring incorrect: n = 4
- Staining incorrect: n = 13

QuIP GmBH, used with permission
QuIP - summary

- 87 participants, 83 results (95%), 60 successful (72%)
- Success possible with CTAs and LDTs
- LDTs: Different clones may be used
  (E1L3N particular success)
- Deviating results: calibration > interpretation
- QuIP: 2nd open ring trial PD-L1 in NSCLC ongoing,
  n=97 participants
LDT-calibration with cell-lines

- TMAs of PD-L1\(^+\) cell-lines: useful for LDT-calibration
- Here: Reference (22C3), CTA-analogue (QC+) divergent (QC-)

<table>
<thead>
<tr>
<th>Control-TMA:</th>
<th>A1</th>
<th>A4</th>
<th>B5</th>
<th>C5</th>
</tr>
</thead>
<tbody>
<tr>
<td>22C3 Kit</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>LDA Site 3, QC+</td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td>LDA Site 1, QC-</td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>LDA Site 6, QC-</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Scheel et al. submitted
Summary

Internal ring trials

- Reproducible PD-L1 scoring and staining possible
- Assays: Reliable;
  \[22C3 \approx 28-8 \approx SP263; \neq SP142\]
- LDTs: LDTs \(\approx 22C3/28-8\) possible;
  6/11 protocols successful;
  calibration essential

Open ring trial

- 60/83 successful (72%)
- Success possible with CTA or LDT
- LDTs: Different antibodies possible
Thank you for your attention!