UK-NCRI Interim PET study

George Mikhaeel, Michael O’Doherty & Sally Barrington
Blinded evaluation of prognostic value of FDG-PET after 2 cycles of chemotherapy in Diffuse Large B-cell Non-Hodgkin’s Lymphoma

Short title: PET after 2 cycles

A sub-study of the R-CHOP-21 v R-CHOP-14 trial

Chief Investigator: George Mikhaeel
To use PET to change treatment (in a future RCT)

We need data on exact prognosis from:

- Homogenous patient group stratified by IPI
- Same Histology e.g. DLBCL
- Same treatment
- Rituximab
- Same criteria for response assessment and change of treatment
- No change of treatment on the basis of PET
- QA in PET centres + Central review of PET
Randomise

R-CHOP 21x8

R-CHOP 14x6

CT > 4 cycles

Baseline PET

Repeat PET > 2 cycles

• Blinded reporting after completion of treatment
  • SUV measurement

Response

No Response

Continue

Off-study
Inclusion criteria

• Age ≥ 18 years.
• Histologically proven DLBCL (central review)
• Bulky stage IA (>10cm) IB, II, III & IV.
• WHO PS: 0-2. Life expectancy >3 months.
• Adequate marrow, kidney, liver and cardiac function.
• Written informed consent
• +ve Baseline PET
Study Design

• Scanning:
  All patients have 2 FDG-PET scans:
  – pre-treatment
  – >2 cycles

Blinding:
• Post cycle 2 scans are archived centrally & treating clinicians are blinded to the scans’ findings
• Nuclear Medicine physicians are blinded to the outcome of treatment
Study Design

Treatment:
- All patients are treated with R-CHOP according to protocol.
- Response is assessed with a CT scan >4 cycles according to IWC criteria

Reporting & Analysis:
- The PET scans are reported in batches after completion of treatment.
- Final Analysis will be performed after completion of recruitment
PET scanning

- QC completed and passed by reference centre
- Reliability of SUV measurement after transfer
- Standard scanning protocol
- Week before 3rd cycle
- 90 min
- Anonymisation
- Central reporting
End Points

Primary Outcome Measure:
• Failure free survival at 2 years

Secondary Outcome Measures:
• Complete response rate
• Overall survival
Statistics

- Assuming that about 50% of patients will have a negative PET scan after 2 cycles and to detect 25% in FFS at 2-years between PET negative & positive groups, with 5% type I error and 90% power, **200 patients** will be required.

Details of Calculation for 25% difference:
- 2y FFS for PET -/+ of 80%/55%: events needed=47, patients needed=191
- 2y FFS for PET -/+ of 75%/50%: events needed=60, patients needed=209
Recruitment

• Target: 200 patients

• March 2010: 142 pts (21 excluded) = 121

• Expected completion: Early 2011
Results

• 97 patients who completed all treatments were analysed

• No outcome analysis

• Comparison of different scoring systems
## PET scoring

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>1 complete disappearance of all abnormal uptake</td>
</tr>
<tr>
<td>Positive</td>
<td>2a MRU Disappearance of most abnormal uptake, but residual low-grade uptake in sites of previous disease, just above the background activity</td>
</tr>
<tr>
<td></td>
<td>2b Partial response Reduction in the abnormal uptake, but significant residual activity</td>
</tr>
<tr>
<td></td>
<td>2c Stable No significant change</td>
</tr>
<tr>
<td></td>
<td>2d Progression Increase in abnormal uptake &amp;/or appearance of new sites</td>
</tr>
</tbody>
</table>
Deauville 5 point Scoring System

• Score 1 (CR): no uptake

• Score 2: uptake $\leq$ mediastinum

• Score 3: uptake $>$ mediastinum but $\leq$ liver

• Score 4: uptake $>$ liver

• Score 5: markedly increased uptake AND new lesion(s) likely to be lymphoma
Comparison of Deauville and R-CHOP substudy scores

<table>
<thead>
<tr>
<th>Score</th>
<th>NCRI Study Score</th>
<th>No of Patients</th>
<th>Deauville Score</th>
<th>Score</th>
<th>No of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>24</td>
<td>1</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>2a</td>
<td></td>
<td>21</td>
<td>2</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>2b</td>
<td></td>
<td>49</td>
<td>3</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>2c</td>
<td></td>
<td>3</td>
<td>4</td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>2d</td>
<td></td>
<td>0</td>
<td>5</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>97</td>
<td></td>
<td>97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Correlation of Deauville and R-CHOP substudy scores

<table>
<thead>
<tr>
<th>Deauville Score</th>
<th>Substudy score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>No of patients</td>
</tr>
<tr>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>97</td>
</tr>
</tbody>
</table>
Comment

- Very few have stable disease (3/97)
- Deauville score may be better in separating significant residual uptake group
### Comparison of Deauville score & Quantitative criteria

<table>
<thead>
<tr>
<th>Deauville Score</th>
<th>SUV$_{\text{max}}$ reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&gt;66%</td>
</tr>
<tr>
<td>Score</td>
<td>No of Patients</td>
</tr>
<tr>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>97</td>
<td>83</td>
</tr>
</tbody>
</table>

Range (67-92%)
Baseline max SUV = 21.4
SUV after 2 x R-CHOP = 6.0
SUV reduction 72% BUT Deauville score 4
Comment

• Good Concordance for scores 1 & 2

• 17/18 of score 3 &
  21/34 (62%) of score 4
  would be responders with >66% SUV reduction

• What predicts response / FFS better:
  – % SUV reduction (regardless of residual uptake)
  – Residual uptake (regardless of initial uptake)
  – ?? Combination

• To improve PPV: cut-off within score 4?
### Cut-offs

<table>
<thead>
<tr>
<th>Deauville</th>
<th>Deauville</th>
<th>SUV reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1+2</td>
<td>3+4+5</td>
<td>1+2+3, 4+5</td>
</tr>
<tr>
<td>45 (46%)</td>
<td>52 (54%)</td>
<td>63 (65%), 34 (35%)</td>
</tr>
</tbody>
</table>

**Substudy**

- Mikhaeel: 41% -ve, 16% MRU, 43% +ve
- Haioun: 60% -ve, 40% +ve

Mikhaeel 41% -ve, 16% MRU, 43% +ve

Haioun 60% -ve, 40% +ve
Conclusion

• Current cohort shows different separation of groups by Quantitative vs 5 point SS

• Final outcome analysis will aim to define cut-off:
  – Best separation of curves (highest accuracy)
  Or
  – Acceptable PPV to use in escalation studies

• Cut-off for interventional studies: may prove to be specific to: disease, treatment, scanning timing, QA / QC of PET