We selected two archetypal examples of distinctly non-PH survival curves for OS:

- Example 1: OS for nivolumab compared with docetaxel for second-line treatment of advanced non-small cell lung cancer (NSCLC),
- Example 2: OS for panobinostat in combination with bortezomib and dexamethasone, versus bortezomib and dexamethasone alone, for the treatment of ipilimumab-refractory melanoma.

The survival curve for nivolumab is shown in Figure 1A, and the survival curve for panobinostat is shown in Figure 1B.

**Objective**

The objective of this work was to test how satisfactorily the scoring methodology of the ASCO-value framework and the ESMO Magnitude of Clinical Benefit Scale captures and represent patient benefit when the PH assumption is not valid.

**Methods**

We identified oncology drugs from the published list of FDA-approved drugs from January 2011 to December 2016. We reviewed published Kaplan-Meier curves and associated statistical summaries for survival endpoints from pivotal trials supporting regulatory approvals.

We selected two archetypal examples of distinctly non-PH survival curves for OS:

- Example 1: OS for nivolumab compared with docetaxel for second-line treatment of advanced non-small cell lung cancer (NSCLC),
- Example 2: OS for panobinostat in combination with bortezomib and dexamethasone, versus bortezomib and dexamethasone alone, for the treatment of ipilimumab-refractory melanoma.

The survival curve for nivolumab is shown in Figure 1A, and the survival curve for panobinostat is shown in Figure 1B.

**Results**

**Table 1: Overview of drug-related characteristics**

<table>
<thead>
<tr>
<th>Drug/Combination</th>
<th>Treatment</th>
<th>Median Survival (months)</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nivolumab</td>
<td>OS</td>
<td>12.2 vs 8.1</td>
<td>0.63 (0.52–0.76)</td>
</tr>
<tr>
<td>Panobinostat</td>
<td>OS</td>
<td>12.0 vs 8.1</td>
<td>0.63 (0.52–0.76)</td>
</tr>
</tbody>
</table>

**ASCOS scores**

- **Panobinostat OS**: Panobinostat achieved a clinical benefit score of 13 based on the primary outcome measure of OS.
- **Nivolumab OS**: Nivolumab achieved a preliminary MCBS grade of 4, based on median OS ≥1 year for the comparator group and a HR ≤0.7. This score is driven entirely by HR reflecting average benefit and captures the value of nivolumab for patients who can expect survival: 12.0 vs 8.1 months, HR (0.63, 95% CI 0.52–0.76).

**Pembrolizumab PFS**

- If the ASCO score was based on PFS data alone (i.e., assuming OS data were reported), pembrolizumab 2mg/kg would have achieved a clinical benefit score of 14 if the HR threshold bonus was added to the score would be 50.
- However, since OS (co-primary endpoint), the score achieved was 14, with HR (0.86, 95% CI 0.67-1.10).

- The score does not appropriately differentiate the clinical value based on the survival: 12.0 vs 8.1 months, HR (0.63, 95% CI 0.52–0.76).
- The survival curve suggests there are some patients who progress early and gain little incremental benefit from treatment with pembrolizumab.

- The plateau that develops in the pembrolizumab PFS curve at around 7 months indicates that patients who progress late may not derive substantial incremental PFS benefit from treatment.

**Conclusion**

- In many regulatory or real-world settings, measures of survival or value frameworks that reduce survival benefit to single scores will not capture important differences in the magnitude of benefit between patient subgroups and can also change markedly with changing survival data.

- The ESMO framework appears to offer greater flexibility, relying on the combination of several summary measures to determine scores; however, hazard ratios are not proportional for OS, the ESMO framework can be driven by an overall HR gain and median survival. However, analyses at a single point in time that are not representative of the survival experience of large proportions of patients.

- In contrast, the ASCO framework consistently prioritizes the HR if it is available, capturing an average patient experience that cannot be simply interpreted and does not represent average treatment effect in the absence of PH.

- The ASCO tail-of-curve bonus compensates in cases where there is evidence of long-term survival benefit that has a limited impact on HR, due to a small number of observations of long survival duration where data are immature. However, it is applied based on passing a threshold (not graduated) and can have a marked impact on total scores.

- Decision makers should be aware of these limitations of single summary measures in modifying treatment decisions - value assessment would benefit from more comprehensive and flexible methods that are able to capture the idiosynsarcies of the complex relationships between survival functions that occur with PH.