Practical Considerations in the Application of Statistical Methods for Treatment Switching

Claire Watkins
AstraZeneca, Macclesfield, Cheshire, UK

Abstract

Treatment switching occurs in a clinical trial when control arm patients switch to experimental therapy during the study. This often happens in oncology trials where patients switch following disease progression, and can result in a large survival difference. An estimate of the survival effect without switching may be required for economic modelling, and several methods have been developed to estimate this. In 2014, the NICE Decision Support Unit published Technical Support Document (TSD) 161 to provide guidance on this.

There are several practical considerations for the statistician or analyst wanting to apply these methods to clinical trials data. The analysis framework proposed in TSD16 is useful, but it can be difficult to apply retrospectively unless the trial was designed with this objective in mind. So are there any trial design features that should be included in the protocol at the start? The raw data must be collected to enable the methods to be applied – what data is that, is it practical to collect it all, what should be done if not? Each method has strong underlying assumptions such as a constant treatment effect or no unmeasured confounders – how could those assumptions be assessed for viability? Several recent health technology assessments have tried to apply these methods and shown either very different results from different models, or have struggled to fit the models at all. Why might that be? What could the analyst do in this situation?

Guidance will be provided on these issues based on experience of applying these methods to real-life data and a review of recent health technology assessments.

Trial design features for the protocol

• Define objective criteria for switch eligibility
• Define switch treatment (drugs, doses) but allow flexibility for emerging therapies
• State intent to statistically adjust for switch and likely primary and supportive methods to justify additional data collection, but allow flexibility to adapt methods to data pattern

Collecting the right data

Table 1: Data collection requirements for switch adjustment methods recommended in NICE TSD16

<table>
<thead>
<tr>
<th>Data</th>
<th>RPSFTM</th>
<th>IPCW</th>
<th>2-stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date of starting switch treatment</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Date of stopping switch treatment</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>All baseline covariates that may influence switch decision or OS</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>All time varying covariates that may influence switch decision or OS, collected until switch or death/censoring</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>All time varying covariates that may influence switch decision or OS, collected until secondary baseline</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Date of secondary baseline (usually disease progression)</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
</tbody>
</table>

What if full data collection is not practical?

This may happen if the decision to adjust for switch is post-hoc, or (for IPCW covariates) if post-progression survival is very long.

• Impute missing data but be clear about assumptions – adds uncertainty
• Summarize impact of missing data methods
• Avoid anti-conservative methods
• Use switch methods less affected by missing data if assumptions reasonable (e.g. RPSFTM)
• Post-progression, collect a few key covariates at less regular intervals than the ideal – preferable to no data due to patient over-burden and dropout

Assessing switch method assumptions

Use assessments that minimise selection bias.

Table 2: Characteristics of NICE TAs with treatment switch adjustments

<table>
<thead>
<tr>
<th>Most common disease area</th>
<th>Renal cancer (3/9 TAs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most common manufacturers</td>
<td>GSK, Roche, Pfizer (2/3 TAs each)</td>
</tr>
<tr>
<td>Most common Evidence Review Group</td>
<td>Pan TAG for early TAs, Liverpool for recent TAs</td>
</tr>
<tr>
<td>Postive recommendation*</td>
<td>6/9 TAs</td>
</tr>
<tr>
<td>OS as (co)primary endpoint</td>
<td>2/10 trials</td>
</tr>
<tr>
<td>Active control</td>
<td>6/10 trials</td>
</tr>
<tr>
<td>Subgroups of longer trials</td>
<td>2/10 trials</td>
</tr>
<tr>
<td>Statistically significant ITT result</td>
<td>3/10 trials</td>
</tr>
</tbody>
</table>

* Negative recommendations all had ICERs (calculated using the committee’s preferred method) of ≥£50,000 aNY. No switch adjustment method was universally preferred

Key findings (Figure 1)

• Confidence intervals wide and usually included the ITT point estimate
• Complex method results (RPSFTM, IPCW, external data) always better than ITT, naive methods (censoring, exclusion) better or worse
• RPSFTM often had smallest HR where used
• If ITT not statistically significant, adjusted result usually also non-significant (in 13/17 cases)
• More inconsistency in results for smaller trials

What to do if results are inconsistent?

Different HRs from the various adjustment methods can lead to very different ICERs.
• Complex methods are preferable to naive methods (less selection bias)
• Check adjusted control arm vs historical data
• Seek medical opinion on plausibility of result and assumptions of methods
• If still no clear candidate, pick the most conservative or middle estimate

References

2. Watkins CL et al. Pharmaceut Statist 2013, 12 348-357
4. Hinks D. Current Medical Research and Opinion 2013, 29 1441-1447

Figure 1: Forest plot of ITT and switch adjusted OS from NICE TAs, sorted by switching rate

• Avoid anti-conservative methods
• Simulate impact of missing data methods
• Compare RPSFTM effect to that from second stage of 2-stage method
• Match control arm switchers to experimental arm using IPC weights and compare OS
• Compare efficacy by disease-stage at entry
• Clinical opinion

[Diagram of forest plot with hazard ratios and confidence intervals]