Pooling Clinical Data: Key points and Pitfalls

October 16, 2012 – Phuse 2012 conference, Budapest
Florence Buchheit
Introduction

- Are there any pre-defined **rules** to pool clinical data?
- Are there any pre-defined **steps** to pool clinical data?
- What is **the right way** to pool clinical data?
Agenda

- Purpose of the pool
- Planning
- Execution
- Possible pitfalls
Purpose of the pool

- showing safety profile across whole program
- improving the precision of incidence estimate for main efficacy analysis
- identifying rare safety signals
- exploration of possible drug-demographic, drug-disease or drug-drug interactions in subgroups of populations.

But

The quality of the pool is dependent on the individual trials used
It is often most appropriate to combine data from studies that are of similar design (dose, duration, control and population)
Purpose of the pool

- Business reasons:
  - Health authorities:
    - Submission of new compound
    - Safety update
    - Further submission
  - Internal use:
    - Exploratory analysis
    - Decision
  - External expert
  - Publications
  - Data Monitoring committee
Planning

- Earlier the better
- Consideration of health authorities, medical expert input
- Pooling strategy: stand-alone versus periodic update

**Stand alone pool**
Example: Publications on specific efficacy parameter for 3 studies
- Study 1 3 4 → June 2013

**Ongoing pool**
Example:
- 1st submission
- Further submissions
- Safety update
- Study 1 2 3 → March 2012
- Study 4 5 → December 2012
- Study 6 7 8 → September 2013
Planning

1. Pooling strategy
   ➡️ Purpose, needs, periodic update

2. Pooling plan
   ➡️ Details of studies, populations, data to include

3. Analysis plan
   ➡️ Details of algorithms, list & shells of outputs

4. Programming specifications
   ➡️ Including mapping data, derived variables, validation plan

5. Pooled analysis datasets creation & validation
   ➡️ Timelines can be split in sub-timelines

6. Pooled outputs creation & validation
   ➡️ Timelines can be split in sub-timelines

7. Dry runs
   ➡️ If agreed

8. Final delivery
   ➡️ Timelines can be split in sub-timelines
Planning

- **Pooling content - safety:**
  - Duration of treatment exposure
  - Adverse events, SAEs (including deaths) and ADRs
  - Specific defined risks of interest
  - Lab values
  - Further areas of potential interest: Vital signs, specifically solicited events captured outside of standard panels (e.g. immunogenicity problems), ECG...

- **Pooling content - efficacy:**
  - Depends on the compound
  - Focus on key endpoint
  - Analysis of subgroup
Execution

- Mapping study level data to pool level:
  - Mapping process: comparison of
    - Meta data
    - Code lists
    - Endpoints
    - Algorithms
  - Increase of mapping work if legacy data, outsourced studies or locally run studies
  - Creation of reports to compare raw and analysis data and meta data for a given study with the final pooled dataset
Execution: raw versus derived

Most common approaches used

- STUDY 1 RAW
- STUDY 2 RAW
- STUDY 3 RAW

+ consistency in derivations across studies

POOL

STUDY 1 RAW
STUDY 2 RAW
STUDY 3 RAW

STUDY 1 DERIVED
STUDY 2 DERIVED
STUDY 3 DERIVED

+ consistency in derivations with study level
+ reduce work in re-deriving endpoints

POOL

Less used – specific cases

STUDY 1 RAW
STUDY 2 RAW
STUDY 3 RAW

STUDY 1 DERIVED
STUDY 2 DERIVED

+ availability of data

POOL
Execution: programs strategy

One hit approach:
- Dataset Study 1
- Dataset Study 2
- Dataset Study 3

- PROGRAM
  - Data standardization
  - Derivations

Dataset pooled

Modular approach:
- Dataset Study 1
- Dataset Study 2
- Dataset Study 3

- PROGRAM
  - Data standardization
  - Derivations

Dataset pooled

+ fewer programs
+ common code to process data across studies

+ more flexible for resources studies and derivations
+ allows preparing studies as they are available
Documentation and validation consideration

- **Documentation:**
  Tracking information regarding data pools centrally to provide an easy reference in cases where multiple requests come.

<table>
<thead>
<tr>
<th>Project Data Pool Tracking Sheet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project:</td>
</tr>
<tr>
<td>CABC123X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Description of data pool</th>
<th>Type of data pooled</th>
<th>Used to support (E.g. 2010 submission, HAQ, HE &amp; OR) etc.</th>
<th>Studies Pooled</th>
<th>Description of data included (E.g. all Safety parameters, All efficacy to 6 months of treatment)</th>
</tr>
</thead>
</table>

- **Validation:**
  - Risks assessment
  - Validation strategy
  - Review data is mapped as required
  - Cross check between pooled and study level data
## Possible pitfalls

### Mapping of treatments:
- Treatments are presented differently in pooled outputs compared to study level
- Special care for studies with multiple periods, down-up titrations, add-on therapies

<table>
<thead>
<tr>
<th>Pooled Trt var</th>
<th>Trt description</th>
<th>STUDIES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Study 1</td>
</tr>
<tr>
<td>T1</td>
<td>Treatment A</td>
<td>A</td>
</tr>
<tr>
<td>T2</td>
<td>Treatment B</td>
<td>C</td>
</tr>
<tr>
<td>T3</td>
<td>Comparator C</td>
<td>A</td>
</tr>
<tr>
<td>T4</td>
<td>Comparator D</td>
<td>C</td>
</tr>
<tr>
<td>T5</td>
<td>Placebo</td>
<td>B</td>
</tr>
</tbody>
</table>

### In derived dataset

### In outputs

<table>
<thead>
<tr>
<th>Output type 1</th>
<th>Pooled trt var</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Trt</td>
<td>T1, T2</td>
</tr>
<tr>
<td>All Active Comp</td>
<td>T3, T4</td>
</tr>
<tr>
<td>Placebo</td>
<td>T5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output type 2</th>
<th>Pooled trt var</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment A</td>
<td>T1</td>
</tr>
<tr>
<td>Treatment B</td>
<td>T2</td>
</tr>
<tr>
<td>All Comp</td>
<td>T3, T4, T5</td>
</tr>
</tbody>
</table>
Possible pitfalls

- Mapping of visit window and timepoints:
  - Are pre-dose assessments in all pooled studies consistent?
  - Are key reporting visits consistent across trials?

- Extension studies:
  - On which treatment should patients who switch treatment from core to extension be considered?
  - Should we consider baseline as core baseline or end of core study?
  - How to handle of duplicate records (events starting in a core study and continuing to its extension)
Possible pitfalls

- Interim analysis
  - How to determine the cut-off point for interim analysis (include all data up to a specific date? a specific visit? include only patients who completed the study? etc.)

- Coding
  - Studies coded using different versions of a coding dictionary (MedDRA and WHO DRL dictionaries).
  - Re-code using the latest version available
Conclusion

- No obvious right or wrong way to pool data but they are many points that need to be well thought in advance.
- Early planning
- Clear pooling strategy
- Programming strategies: raw versus derived / one hit approach versus modular approach should be considered depending on the pool objectives, update, data availability...
- Pooling will be facilitating if adherence to standards
Questions ?