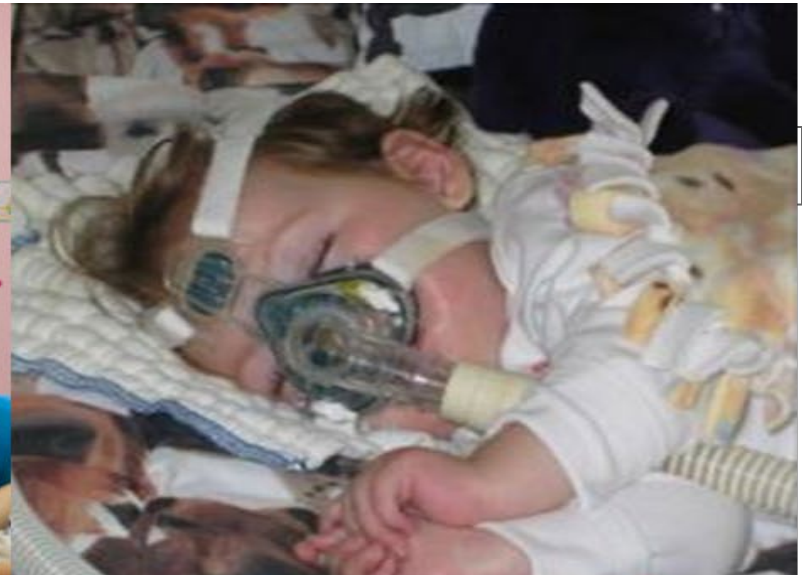
When a Threshold Crossing approach may and may not be appropriate: A Case Study in SMA

EFSPI Regulatory Statistics Workshop, 24-25th September 2018
Carol Reid and Uli Burger





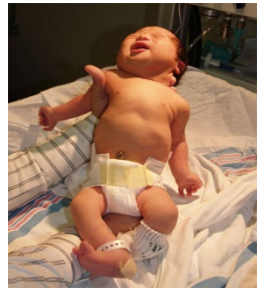
Outline

- Introduction to Spinal Muscular Atrophy (SMA)
- Designing a Study for Patients with Type 1 SMA
  - Threshold Crossing Approach
  - Primary Endpoint
  - Secondary Endpoints
- Designing a Study for Patients with Type 2 & 3 SMA
- Conclusions

Introduction to Spinal Muscular Atrophy (SMA)

- Genetic, progressive motor neurone disease characterized by muscle atrophy and weakness
- Continuous spectrum of symptoms
  - Patients classified into types according to highest level of function achieved

Type	Severity	Age of onset	Typical symptoms	Lifespan
I	Severe	0-6 months	• **Never sit** • Respiratory failure	< 2 years
II	Intermediate	7-18 months	• Sit, **never stand** • Respiratory complications likely • Wheelchair-bound	> 2 years
III	Mild	> 18 months	• Walk at least once in lifetime • Muscle weakness	Adult
IV	Mildest	2nd and 3rd decade	• Gradual weakening of muscles in adulthood	Adult

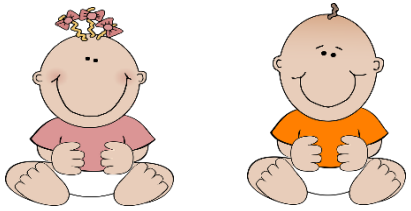


- First treatment for SMA approved in Dec 2016 (FDA), 2017 (EMA)

Designing a Study for Type 1 SMA

- Rare disease with no approved treatments
- Study in rapidly declining infants with short life expectancy
  - Placebo control ethical?
  - Strong push from patient groups against randomized trials
- Patient population clearly defined: Genetic confirmation of SMA + clinical symptoms
- Natural history of the disease is well defined
  - Type 1 SMA infants never achieve the motor milestone of sitting
- Primary endpoint: Proportion of infants sitting without support at 12 months of treatment
  - Video-recorded and centrally assessed by 2 independent raters
  - Objective & clinically meaningful.
- Single arm study selected

  Threshold crossing approach to assess success



Threshold Crossing (Eichler et al. 2016)

- Upfront definition of an appropriate estimand defining
  - The treatment-eligible population
  - The variable of interest
  - The measure of intervention effect
- Counterfactual determined from existing RWD and/or past RCT data
- Efficacy threshold clearly higher (high bar) than the estimate of the counterfactual is set and agreed by relevant decision-makers alongside a detailed study protocol and analysis plan
- Patients who meet the predefined treatment eligibility criteria are enrolled in a single arm study
- If the threshold is crossed, efficacy is judged to be established
  - A negative result would lead to an additional study or termination of product development

"Threshold Crossing": A Useful Way to Establish the Counterfactual in Clinical Trials? Eichler, HG et al., Clinical Pharmacology & Therapeutics, 2016

Primary Endpoint Threshold & Analysis

- Type 1 infants never sit
- Threshold to be statistically differentiated set at 5%
- Hypothesis to be tested is proportion of infants who sit at month 12 (p)

  Ho: $p \leq 5\%$ vs Ha: $p > 5\%$
- Tested using exact binomial test
  - Reject Ho if one sided p-value $\leq 5\%$
  - Study will be positive if lower limit of two-sided 90% Clopper-Pearson (Exact) confidence interval is above the threshold of 5% set for the primary objective
- Note study will be positive as soon as minimal number of infants sitting needed to achieve positive test is observed
- Study will continue for an unbiased estimate of the primary endpoint and for the assessment of secondary endpoints

Secondary Endpoints

- Secondary endpoints included to assess progression of other important aspects of the disease
- Secondary Endpoints include:
  - The proportion of infants who achieve >40 in the CHOP-INTEND
    - 16-item scale measuring motor function (hand-grasping, rolling, head control etc)
  - The proportion of infants who achieve motor milestones (head control, rolling, sitting, standing, walking)
  - Time to death or permanent ventilation
- Natural History is less well-defined
  - Can be variable and dependent on individual clinician practice & family preference
  - Mortality is generally dependent on the aggressiveness of pulmonary intervention
- Where possible, secondary endpoints will be assessed against pre-defined thresholds determined from observational and clinical trial data
- As a sensitivity analysis secondary endpoints will be compared with data from a retrospective chart-review study conducted at the same sites as the pivotal study

Thresholds for Secondary Endpoints

- Determined from similar cohorts of patients constructed from RWD sources and clinical trial data (literature search and databases)
  - All data sources and reasons for exclusion are documented in an appendix to the SAP
- Where patient level data available, summary data generated from patients with similar exclusion/exclusion criteria when possible
- Where only summary data available point estimates and confidence intervals extracted or derived
- When multiple sources of data existed for the endpoint, the cohort with baseline characteristics most similar to the targeted study population were selected
- High bar: Threshold is based on the upper limit of the 90% confidence interval (CI) derived from the historical data
- Thresholds were pre-defined in the SAP prior to study start
- Any new external data will also be presented and may be used for sensitivity analysis

Chart Review Study

- A retrospective chart-review study aims to provide additional data on the natural history of selected secondary endpoints to provide further context for the results
- Infants with the same key inclusion criteria as the pivotal study and matched on key prognostic factors
- Data extraction in the same centres to minimize any centre effects resulting from differences in standard of care
- Start date selected to ensure that the standard of care for infants included is as comparable as possible to current practice
- End date selected to eliminate the risk of over-representing infants ineligible for clinical trials and of informative censoring
- Selection of the most recent eligible chart during the study period to further minimize the impact of any temporal trends in patient outcomes
- Well-defined, deterministic, chart-selection process to eliminate the risk of inappropriate selection of charts

Study Design for Type 2 and 3 SMA

Type	Severity	Age of onset	Typical symptoms	Lifespan
II	Intermediate	7-18 months	• **Sit, never stand** • Respiratory complications likely • Wheelchair-bound	> 2 years
III	Mild	> 18 months	• **Walk at least once in lifetime** • Muscle weakness	Normal

Primary Endpoints focus on **motor function:**

E.g. Motor Function Measure (MFM-32)

32-item scale administered by physiotherapists that evaluates physical function in 3 dimensions:

- standing and transfer, axial and proximal function, distal motor function

- Items scored from 0 (unable) to 3 (fully able)
- Objective Scale (limited scope for subjectivity)

Single Arm or Randomised Controlled Trial?

- Again clear patient preference against randomized trials but still possible to recruit patients into a placebo-controlled trial
- Natural history of the disease is less well understood
  - Patients decline but at variable rates depending on age and other (unknown) factors
  - E.g. Patients aged around 5-15 decline at a greater rate than older patients
- Smaller effects relevant: Small changes (or stabilisation) are meaningful to patients
  - Markedly different changes less likely in older patients
- Possible to set a threshold for success
  - E.g. Lower 95% CI > 0 in a population expected to decline
- May need a longer and/or larger study to be conclusive. Differentiation of smaller and medium size effects is difficult
- Primary endpoint may have some assessment bias

Conclusions

- In Type 1 SMA a single arm study assessed using a threshold crossing approach is appropriate
  - High ethical demand
  - Selected primary endpoint is objective with little assessment bias, clinically meaningful with known natural history. High bar versus natural history
  - Thresholds for some secondary endpoints determined from available natural history and clinical trial data are less clearly defined but still provide useful supportive information
  - Additional information from a chart review study provides supportive data from the same sites for sensitivity analyses
- In type 2/3 SMA a randomized study is more appropriate
  - Potential primary endpoints have limited scope for assessment bias
  - Natural history is less well-defined
  - Smaller effects may be clinically meaningful but cannot be differentiated versus natural history
  - Non-controlled study may need to be larger and/or longer to be convincing, with a potentially unrealistically high bar

Doing now what patients need next