Clinical Trial Design Considerations for Antifungal Development

Aaron Dane

Outline

- Use of external controls to supplement clinical trial data
  - When is this appropriate as opposed to an RCT
  - Points to consider when using an external control
  - Use of external controls alongside single arm trial or an RCT

- Alternative Statistical Criteria in a Study of Rare Molds

Use of External Controls in Limited Populations

Use of external data with small patient numbers

- It may only be feasible to recruit 50-100 patients with rare molds in a reasonable time period

- Choice between a very small RCT or a single arm trial
  - Small RCT gives randomization, but heterogeneity may make comparing treatments difficult
  - Non-randomised study means comparing with externally generated data
  - When patients have no treatment options a single arm study may be the only option

- Randomization is generally preferable
  - But if there is no clear SOC or a robust external dataset is available use of external data may provide more reliable information

Robustness of external controls

- Contemporary and matched controls most useful
  - Due to similarity of disease setting and standard of care
  - How "contemporary" does the external control have to be?

- Data validity - Can we verify the data used?

- Potential for bias or lack of comparability to RCT
  - Are patient populations and treatment of patients similar?
  - Were data collected under similar trial conditions?
  - Are regions, resistance levels … similar?
  - Is the endpoint defined in same way?
  - Are there differences in reporting of cases?
  - Is data available for all external patients or a selected subset?
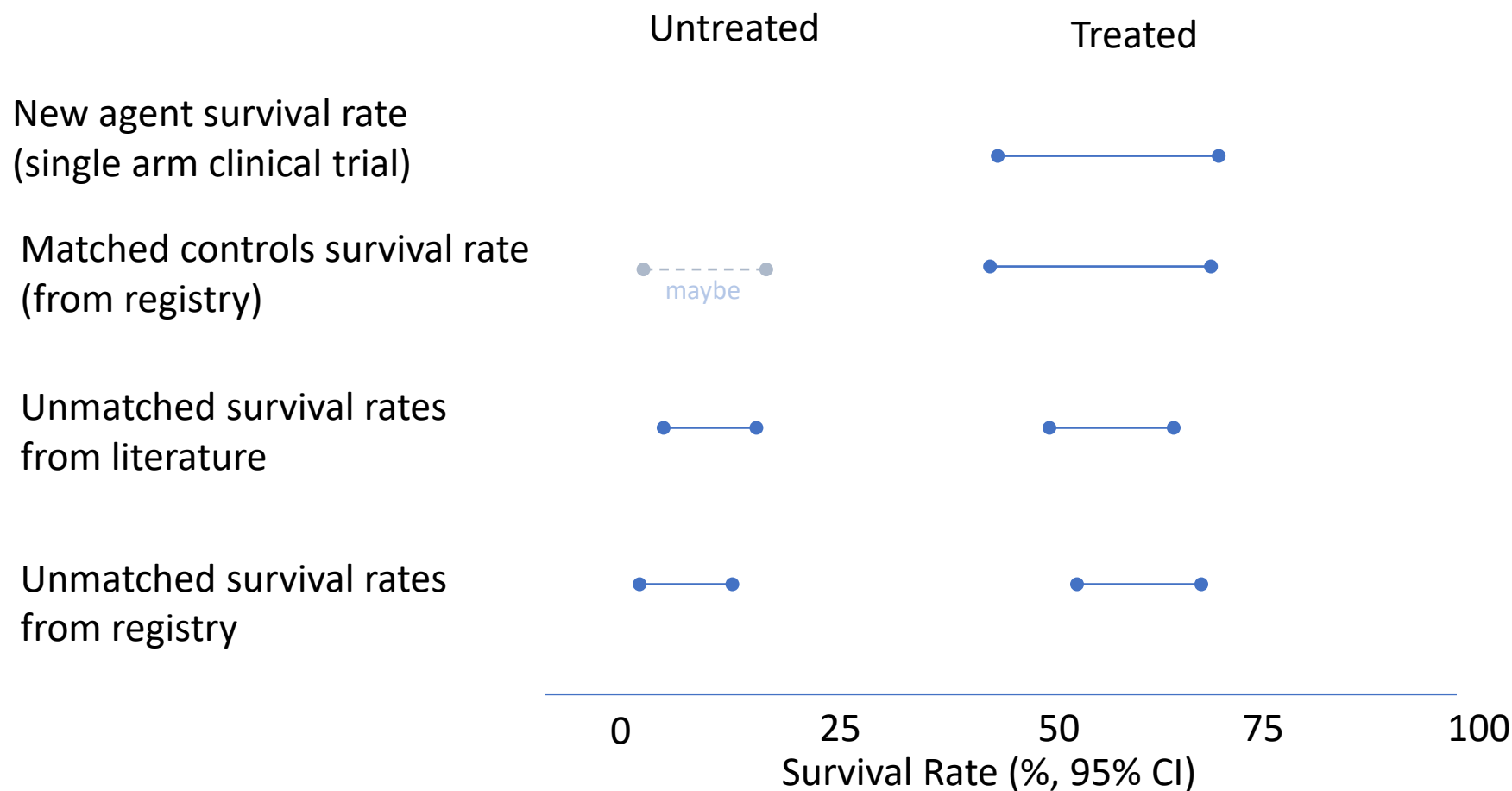 - Is matching possible or necessary?

Patient Population and Patient Care

Are key features similar for external control and clinical trial?

- Are patients identified in same way?
  - Are all available patients with disease in question included in the external cohort, or is this a selected subset?
  - Are external controls and trial patients identified at same point in disease course?

- Is patient prognosis similar?
  - Are risk factors consistent between external cohort and clinical trial?
  - Are risk factors consistent across sites and countries in external cohort?

- Is there a consistent approach to management of patients in external control?
  - Consistent management within a country, and between countries?
  - Is the dose and duration of treatment appropriate?
  - Is the SOC for each country/site sufficient to allow comparison

Demonstration of clear benefit when comparing single arm trial with external data

Based upon approach of Isuvaconazole and FungiScope registry

Using External Data With an RCT

Bayesian-augmented Controls

- As an example, a traditional design may require 700 patients (350/arm)[1]
  - Augmented control design could recruit less than 700 patients in 2:1 ratio
  - Would supplement with data from external clinical trial using same comparator

- Provided control group response rate in clinical trial is similar to external control rate this could allow similar type I error and power, with fewer patients
  - Viele[2] has outlined the possibilities of this approach
  - Exact details are case dependent, but has potential for more efficient trials

- Key risk is when the "true" control arm response rate is not same as external data
  - If true control group rate is lower this reduces power (higher sponsor risk)
  - If true control group rate is higher this increases type I error (regulators risk)
  - This assumption is critical when applying this approach

[1] Typical sample size for trial with 80% response rate and 10% NI margin
[2] Viele, K et al. Use of historical control data for assessing treatment effects in clinical trials. Pharm. Statistics 2013.

Alternative Statistical Criteria for a Randomized Trial for Rare Molds

This approach has been developed in collaboration with Prof. Nigel Stallard (Warwick University, UK), Paul Newell and John Rex

This talks outlines an abbreviation of the original presentation at the FDA-Pew Workshop "Enhancing the Clinical Trial Enterprise for Antibacterial Drug Development in the United States" in November 2019

Decision making in clinical trials

- What are we most interested in for any clinical trial?
  - To be confident we can show an effective treatment works
  - To be confident we will not approve ineffective treatments

- Can we look at this differently for rare molds?
  - These patients are hard to find during clinical trials
  - It is better to provide a framework for evidence of effect in rare molds for physicians rather than having no data
    - without this decisions will be made without data
  - We have looked to draw on ideas used in the orphan drug area

- Even with smaller studies we need a framework for decision making
  - Sponsors need to understand what is required to assess study risk
  - Provides clarity regarding decision making criteria at the design stage

Aim

- To propose a framework for decision making and sample size for rare pathogen studies where feasibility is extremely challenging and when (sub-optimal) therapies are still available

- This is not about performing an interim analysis where we decide to continue to recruit more patients, but rather how to understand the risks with a smaller study

- This talk focuses on traditional frequentist statistics, but has also been considered within a Bayesian framework

Large v Small trials with rare pathogens

- Larger trials lead to higher power
  - But if the trial is too large (or takes too long) this deprives patients of an effective therapy and may mean it is not feasible to develop the drug

- Smaller trials may be more feasible
  - But if trial is too small we have larger chance of making the wrong decision

- Common theme: How to work with the only (small) dataset possible?
  - Can we show that there is a "sweet spot" for sample size?
  - There can be diminishing returns outside the "sweet spot"

Can we define a "sweet spot" to balance these questions

What are we aiming for?

- If TEST is worse than CONTROL…
  - Every patient randomized to TEST risks a worse outcome
  - If TEST is approved, this problem is perpetuated
  - *Mitigation: Within a small trial, avoid incorrect approval (Type 1 errors)*

- If TEST is better than CONTROL…
  - Every patient randomized to CONTROL risks a worse outcome
  - If TEST is <u>not</u> approved, this problem is perpetuated
  - *Mitigation: Within a small trial, keep the power high*

- If TEST is similar to CONTROL…
  - We still want to make additional therapies available
  - *Mitigation: Within a small trial, keep the power high*

When we run the trial we do not know which of these situations is true so we must understand the type I error and power for a range of study sizes

Recap: Finding the "sweet spot"

- We need to find a sample size where:
  - We have a good chance of success when effective
  - We have a low chance of approval when ineffective
  - We have a reasonable chance of success when similar
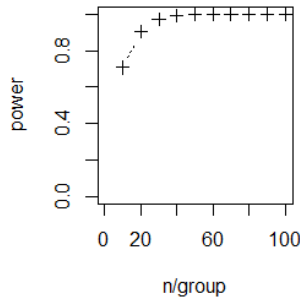 - The expected number of patients benefitting is maximised

The following plots summarize this information…

Using 80% CI could be reasonable with high unmet need
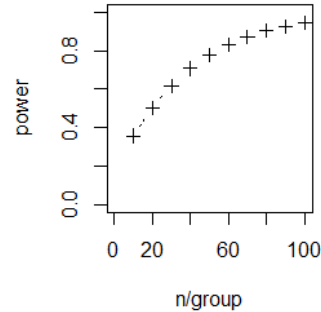
Using 80% CI & 20% NI margin

How often is TEST selected?

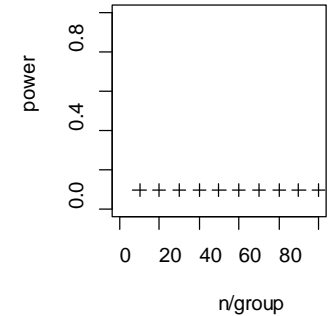**When Effective**
Test is better (60% v 40%)

Power is high for positive effect

When Similar
Test and Control both 40%)

Power is reasonable for 50-60/arm when response is similar to control

When Ineffective
Test is worse (20% v 40%)

There is a low (10%) chance of incorrectly concluding NI When Test is worse

Why use different statistical criteria?

- In addition to the power and risk of incorrect approval, there are additional considerations when running a study in a small population

- In a trial where one treatment is less effective, many patients receive this sub-optimal therapy
  - 50% receive suboptimal therapy with 1:1 randomization
  - In a limited population a large proportion of the overall population with the disease outside of the trial would not receive effective therapy
  - The size trial of trial to maximise the total number of patients outside of the trial expected to benefit from the best therapy can be calculated
  - It is possible to display this graphically but that is beyond the scope of this presentation

- In summary, a much larger study does not always provide the best outcomes in a limited population.

Considering Alternate Statistical Criteria

- This is a framework to display trade-offs when only a small trial is possible
  - What are reasonable false positive and false negative error rates?
  - As a community we need to decide how to trade these risks when we cannot run large trials

- Data on 100 patients with rare mold can be very informative
  - But we need clear criteria for success that can be agreed ahead of trials
  - How to maximise our chances of approving a more effective new drug with (say) 100 patients?
  - How to limit the risk of approving a less effective new drug with (say) 100 patients?

Considerations of power, the chances of an incorrect approval and the overall number of patients benefitting from therapy can be used to agree success criteria in trials of rare molds

Summary

- Studies of rare molds are incredibly challenging to recruit

- It is not possible to design studies in a traditional way with traditional statistical criteria

- External controls can help provide robust evidence when only a small dataset is possible
  - Data need to be robust and comparable to the clinical trial
  - A large treatment effect is important given different data sources

- Alternative statistical criteria can be considered for rare molds where there is a high unmet need

Back-up Slides

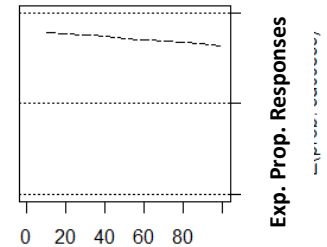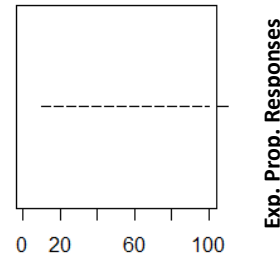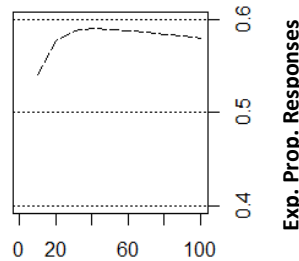# Expected number of patients benefitting from therapy

Using 80% CI & 20% NI margin

When Effective
Test is better (60% v 40%)

When Similar
Test and Control both 40%)

When Ineffective
Test is worse (20% v 40%)

Expected number of patients benefitting during and after the trial

Number of patients responding after the trial is lower if the trial is large and one treatment is less effective

Number of patients benefitting is the same as both treatments have same response

Number of patients responding after the trial is lower if the trial is large and one treatment is less effective

Expected Number of Responses
Looking at the two sample size scenarios side by side

Responses in trial   **Responses following trial**

RCT sample size	RCT success (TEST)	RCT sucess (Control)	Probability TEST is taken forward	If TEST taken forward	If Control taken forward
50/arm	30/50	20/50	0.983	540/900	360/900
200/arm	120/200	80/200	0.999	360/600	240/600

Expected number and proportion of successes in the overall population

50/arm: 30 + 20 + (0.983 x 540 + (0.017 x 360) = 587   →   (i.e. ~59% of all patients)

200/arm: 120 + 80 + (0.999 x 360 + (0.001 x 240) = 560   →   (i.e. 56% of all patients)

TEST responses in RCT

Control responses in RCT

Expected responses after RCT is
Probability TEST taken forward x number responses if TEST taken forward plus
Probability Control taken forward x number responses if Control taken forward

The expected number of responses is important when understanding whether a patient may benefit from the approval of a new drug so is a key element of the following slides