# Statistical Considerations for Antibiotic Drug Development

**Aaron Dane, AstraZeneca**
**Biometrics TA Head (Infection)**

**CTTI Statistics Think Tank, 19 November 2014**

AstraZeneca

# Ideas in this talk

- What is the issue and how could we approach it?
  - The tiered regulatory approach

- What are the options when only smaller RCTs are possible?
  - Statistical criteria
  - Bayesian approaches

- Interpretation of information on small numbers of resistant pathogens
  - So small that any inferential testing is challenging
  - Formal demonstration of superiority is not feasible
  - Use of supplemental information from external sources
  - Issues and methods with using all available information



# Background to studying rare pathogens

- For registration, we traditionally expect
  - Two substantial trials per indication (e.g., two UTI trials)
  - Typical size/trial for antibiotics: ~1,000 patients
- But, what if the target disease includes a less common, but important, pathogen or type of resistance?
- We need to run trials when resistance is less common in order to have treatments available in an epidemic
- When only limited clinical data for these important subsets are possible, programs should consider how to best use all available data



# The Challenges of Superiority

- Superiority trials are preferred when possible, as they provide a clear interpretation of the clinical trial

- Showing superiority on a clinical endpoint is not routinely possible; either:
  - Need to knowingly study ineffective or toxic comparators in seriously ill patients[1], or
  - Formal demonstration of superiority is challenging when the patients of interest are rare due to sample size limitations

- Superiority may be possible for highly resistant pathogens
  - This is the case when standard therapy is ineffective
  - However, new drugs will make such comparators unethical, and any superiority trials infeasible in the future
  - Therefore, non-inferiority approaches still need to be considered



[1]Nambiar et al. Clin Pharm Ther 96:147-149, 2014.

# Why is superiority so difficult in an RCT?

**Recruited Population**

**N=300/arm**

**Confirmed pathogen for primary population**
**(eg, pseudomonas, 3-20%)**

**N=9 to 60/arm**

**Pathogen resistant to all other therapies**[2]

**Low N; also removed from study at day 3-4**

**Plus confounding with co-morbidities**

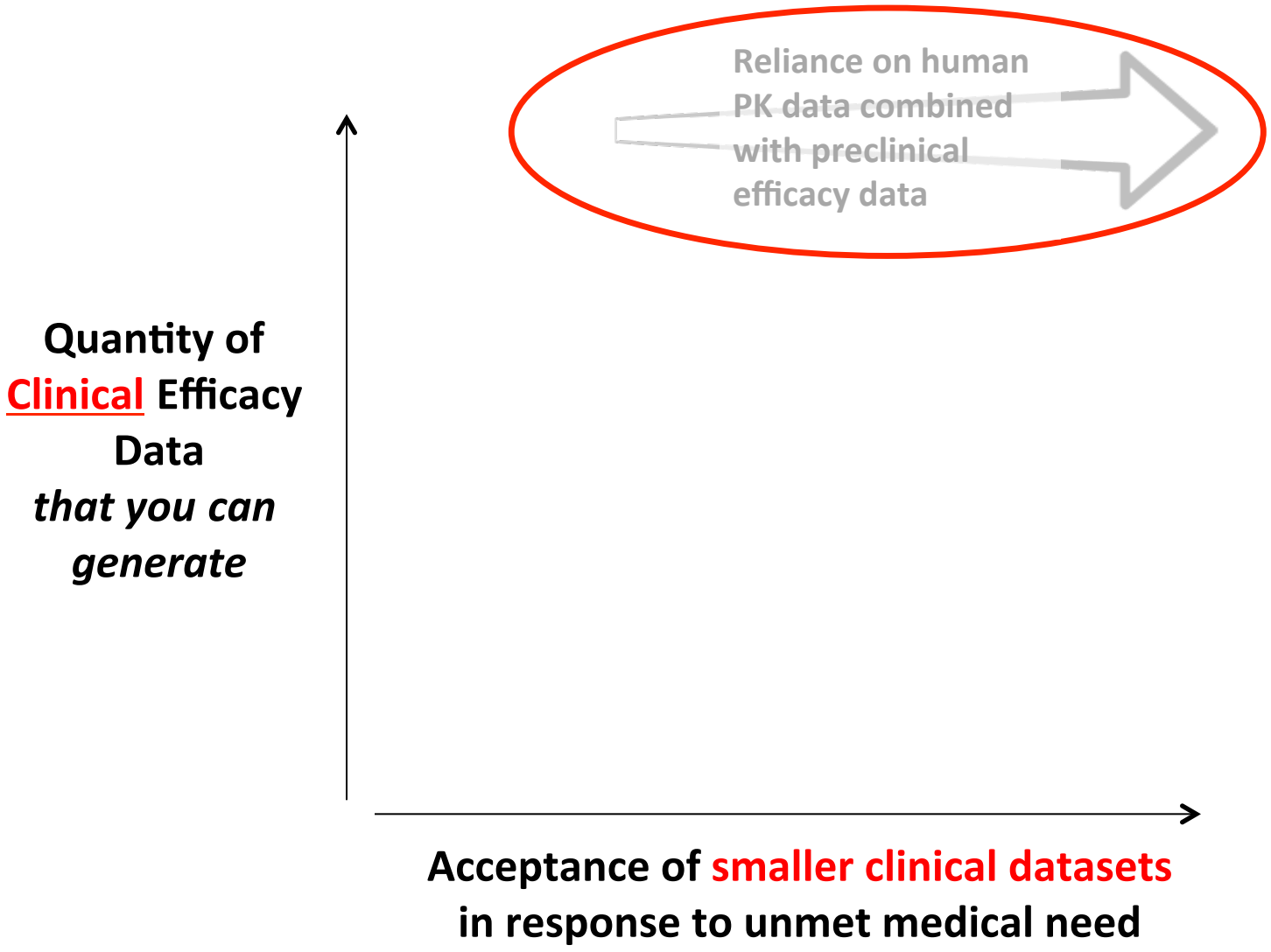
**Formal superiority not feasible, even before other potential confounders**



[2]Sbrana et al. CID 56:697-700, 2013 showed that it is difficult to find 100% resistance, even with challenging pathogens

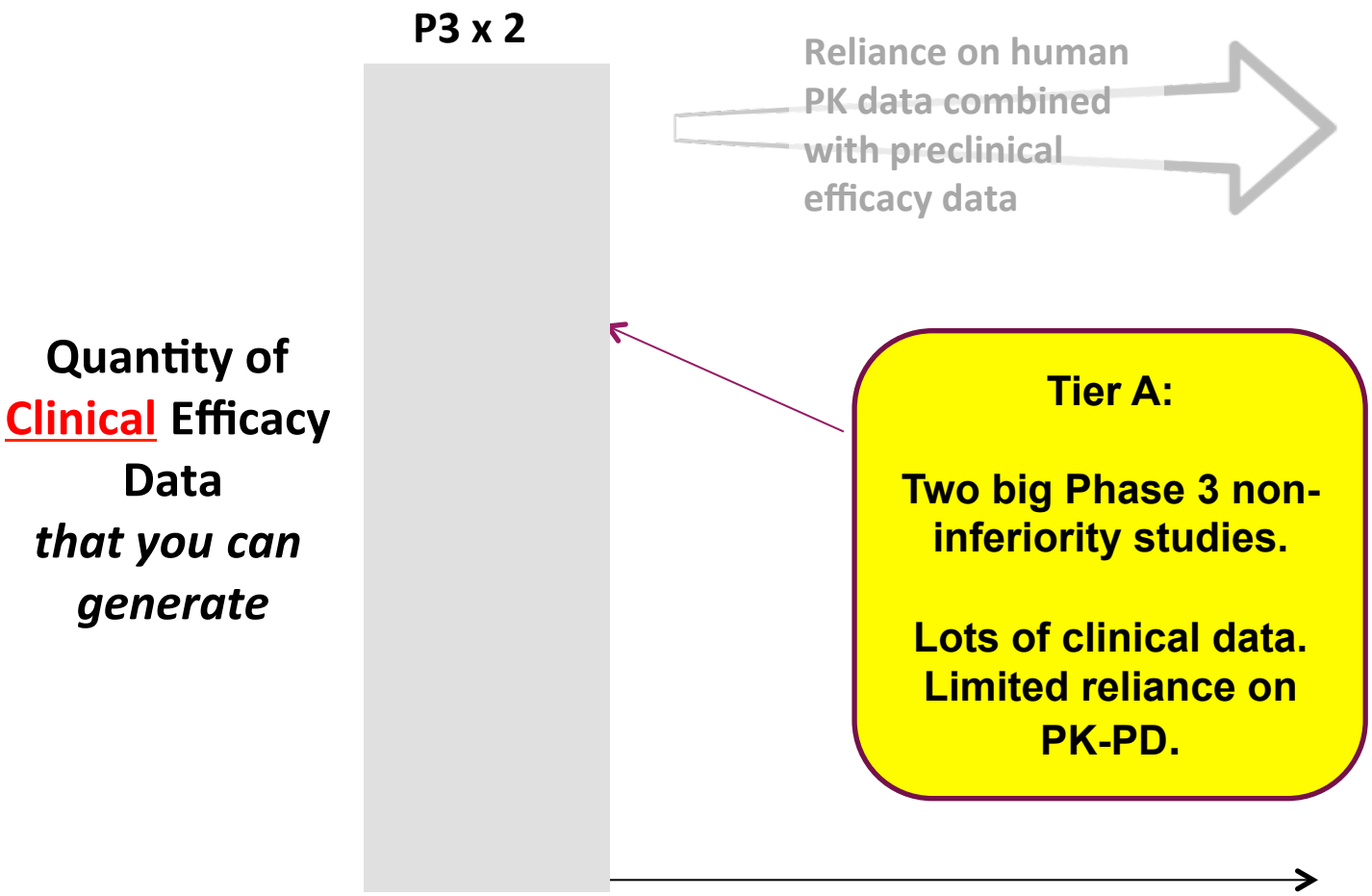
# Development Options as Tiers

Rex et al, Lancet Infectious Diseases, Volume 13, Issue 3, Pages 269 - 275, March 2013



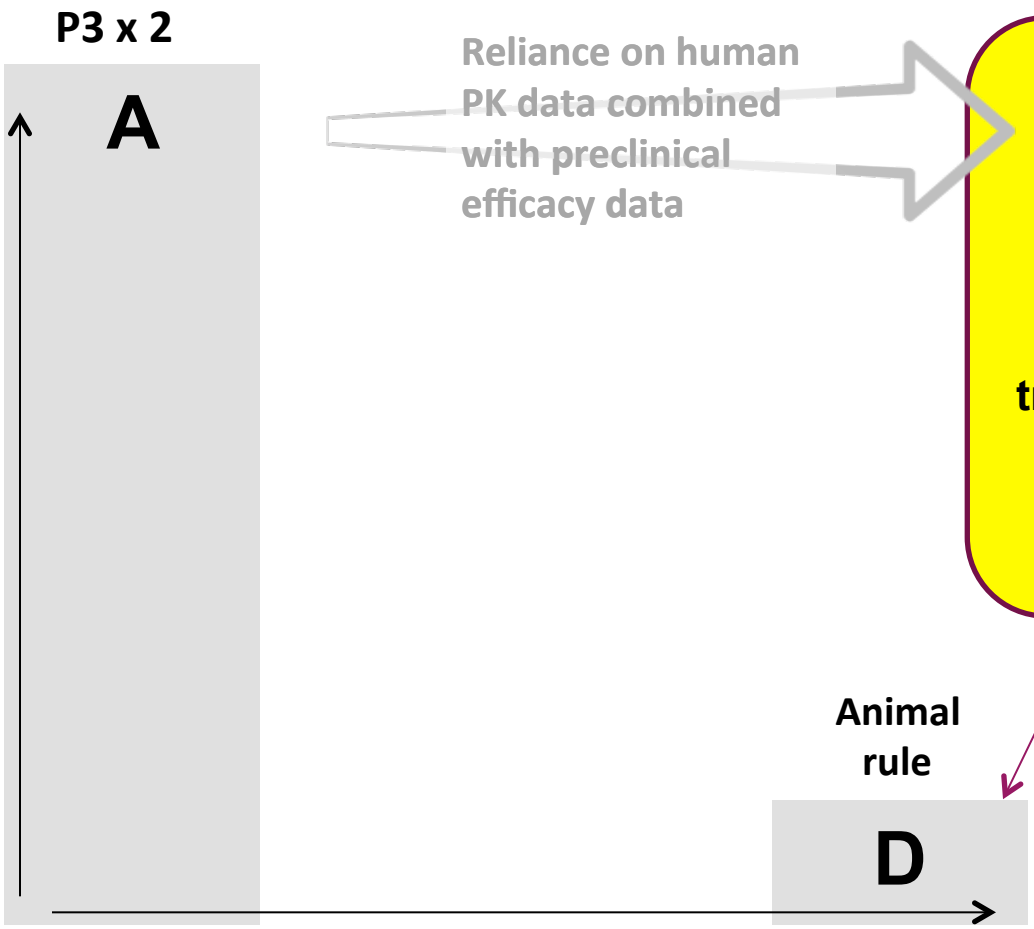
# Development Options as Tiers

## Tier A: The traditional approach



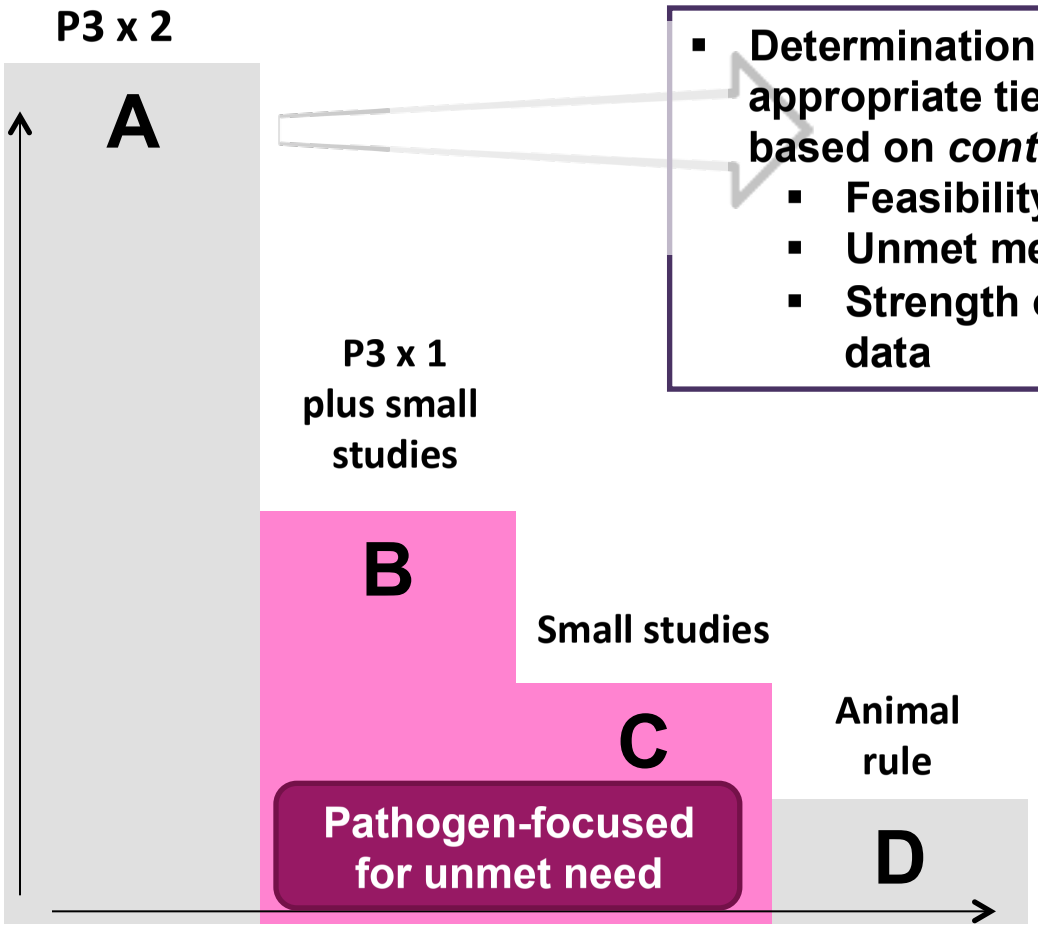
# Development Options as Tiers

## Tier D: The animal rule



# Development Options as Tiers

## Tiers B & C: Pathogen focussed developments




- Determination of the appropriate tier should be based on *context*:
  - Feasibility
  - Unmet medical need
  - Strength of the preclinical data

# Tier C approaches use the "totality of data"

- High unmet need justifies accepting more uncertainty regarding efficacy and safety in product development.
  - Severity of unmet and strength of totality of data agreed with agency at the outset
  - A comprehensive, supportive pre-clinical program is vital
  - The level of uncertainty should be explicitly described and discussed

- Pre-clinical
  - Increased utilization of pre-clinical efficacy & prominent use of PK/PD data in the assessment of new agents
  - Could strength of PK/PD information be considered pivotal information?

- Clinical
  - Conduct small RCT to generate some efficacy and safety data in controlled setting
  - Use safety data from all trials relevant to that product or combination
  - Clear Risk Management Plan appropriate for an area of unmet need

***The following sections will cover situations when (1) traditional RCT sample sizes are not feasible in a reasonable timeframe, and (2) situations when only very small amounts of data are feasible***



# Ideas in this talk

- What is the issue and how could we approach it?

- What are the options when only smaller RCTs are possible?
  - Statistical criteria
  - Bayesian approaches

- Interpretation of information on small numbers of resistant pathogens
  - So small that any inferential testing is challenging
  - Formal demonstration of superiority is not feasible
  - Use of supplemental information from external sources
  - Issues and methods with using all available information



# When only a Small Phase 3 RCT is Possible

- An RCT is a powerful source of unbiased data
  - It addresses safety & efficacy and reduces risk for developer & regulator
  - It would be preferable to produce some RCT data, but less than usual

- Methods using all of the available information or more clearly understanding uncertainty are important



# Different statistical criteria

- What result will support approval?
  - For high(er) unmet need, greater degree of uncertainty may be reasonable

**Options to make adequately powered trials more feasible**

- Wider NI margin
  - Often evidence of big benefit over placebo from historical data
  - A wider margin with less discounting justified in areas of unmet need

- Alternative value of alpha
  - Traditional 2.5% alpha means we have a $\leq$2.5% chance per trial of observing data consistent with NI conclusion if new agent truly worse
  - Applying alpha of 5% or 10% means a 5% (or 10%) chance this occurs



# Different statistical criteria

## Effect of changing margin & alpha

- With typical parameters (80% response, 90% power)
  - Usual alpha = 0.05 (0.025 as one-sided) and 10% margin
  - Size would be 337/arm evaluable patients

- This can be reduced by 2/3rd or more
  - alpha = 0.10 (0.05 as one-sided), 15% margin → 122/arm

***Evaluable patients needed/arm***

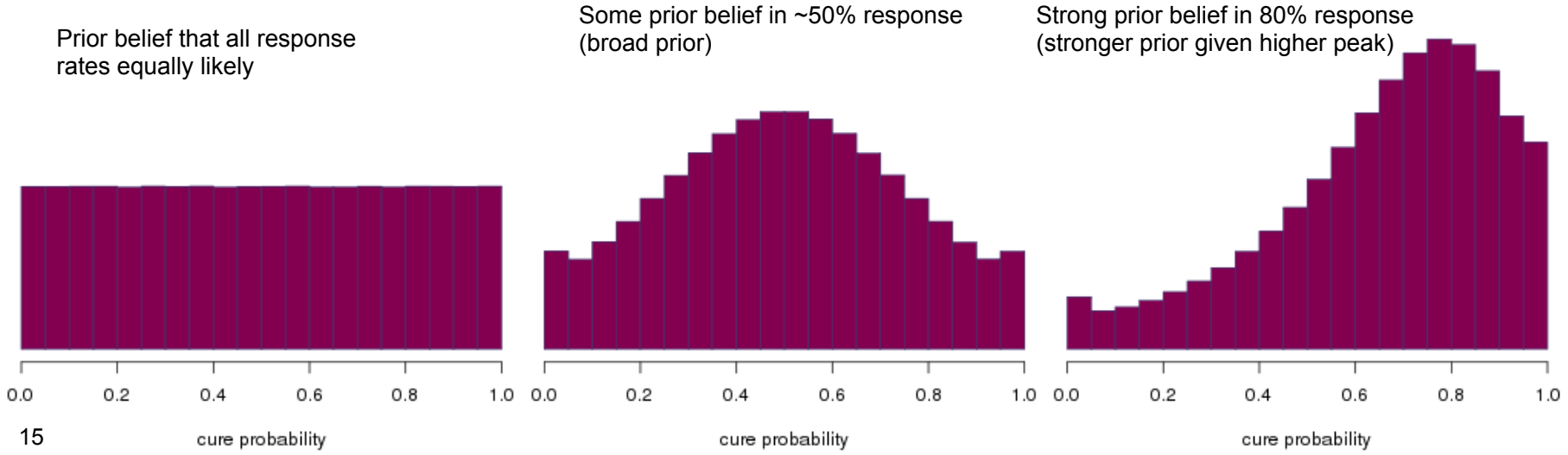
| 1-sided alpha | NI margin | | |
|---|---|---|---|
| | -10% | -15% | -20% |
| 0.025 | 337/arm | 150/arm | 85/arm |
| 0.05 | 275/arm | 122/arm | 69/arm |
| 0.10 | 211/arm | 94/arm | 53/arm |

# Bayesian Approaches

- Frequentist analysis approaches make no prior assumption about the anticipated response of the experimental or control agent

- We generally have more confidence in the expected level of efficacy of experimental or control taken from sources external to the RCT
  - For the experimental arm this can be taken from PK/PD data
  - For control agent, this can be focussed on recent clinical trials

- The possibilities range from making no assumptions regarding expected response to having a strong belief in the expected response depending upon the supportive data available

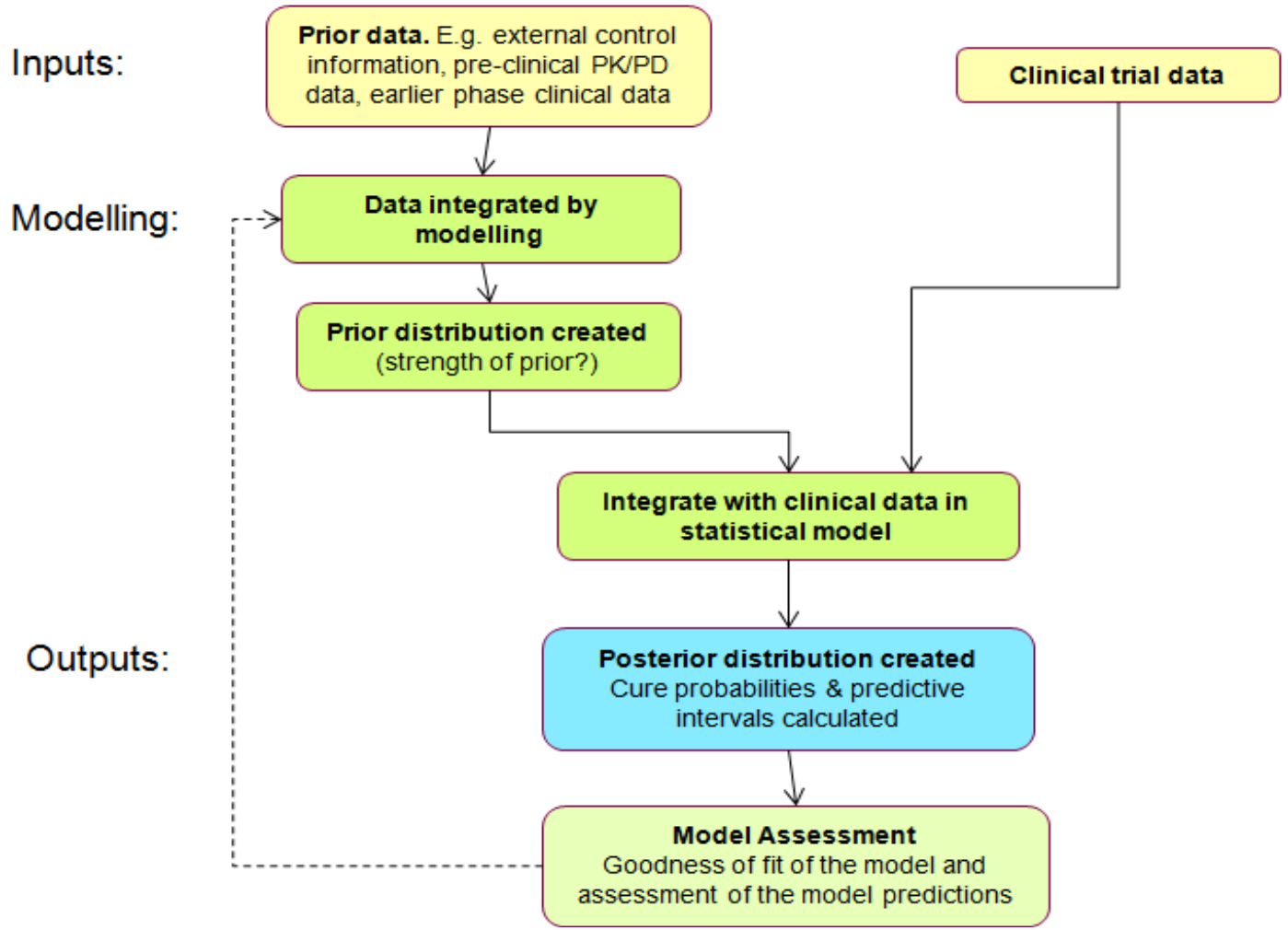





Note: peaks at tails of distribution used to incorporate additional uncertainty in prior belief whilst retaining best estimate of expected response

# Bayesian Approaches

## Role of prior distributions



# Bayesian Approaches

## Bayesian-augmented Controls

- Designs aim to maintain type 1 error and power to traditional designs while maximizing sample size savings via borrowing from historical data
- With dynamic borrowing
  - The amount of borrowing depends on precision among control trials and similarity of historical data to concurrent control
  - Results in a reduction in sample size, more patients on treatment and increased power when true control rates near observed historical data
- Possible risks when RCT control rate differs substantially from observed historical data
  - Inflated type I error when true control rates substantially above observed historical control rates (but still less than static borrowing)
  - Decreased power when true control rates substantially below observed historical data
  - Due to the growing resistance problem, we believe downward drift will be a more likely risk
- As a result, similar clinical setting and patient population is needed, along with a strong belief in similar response rates



# Bayesian-augmented Controls

## Example – Bayesian approach can reduce patient numbers

- **Traditional (fixed) Design, N=750**
  - Operating Characteristics
    - Active control cure rate = 83%
    - 10% NI margin
    - 90% power and 5% two-sided significance
    - 20% dropout; 1:1 randomization
  - 375 patients per treatment (n = 750 total)

- **Bayesian approach, N=600**
  - N=600 with 2:1 randomization and borrowing
    - 400 subjects on treatment
  - Assumes historical control of 83%
  - Working hypothesis
    - NI concluded if 1-sided 97.5% CI for trt effect $\geq$ -10%
    - Type I error: conclude NI when test trt >10% worse
    - Power: ability to correctly conclude NI

**We need to control type I error at <0.025 and retain reasonable levels of power (~90%)**

| True Control Group Rate | Type I error | | Power | |
|---|---|---|---|---|
| | Traditional | Bayesian | Traditional | Bayesian |
| **78.0%** | 0.024 | 0.006 | 84.2% | 70.2% |
| **80.5%** | 0.026 | 0.007 | 87.4% | 86.1% |
| **83.0%** | 0.026 | 0.017 | 91.0% | 94.2% |
| **85.5%** | 0.028 | 0.045 | 92.9% | 96.6% |
| **88.0%** | 0.030 | 0.100 | 95.4% | 97.6% |



Viele et al (2013) ***Adaptive design for a Phase 3 trial of cUTI that utilizes historical control data***, manuscript in final draft

# Bayesian Approaches

## Use of PK/PD data to construct prior distributions

- Preclinical and surveillance data provide relationship between PK parameter of interest (e.g. AUC) and MIC
  - Provide target levels for dosing to achieve microbiological kill.
  - PD target taken from pre-clinical experiments
  - PK estimates taken from human PK data generated in early clinical trials
  - Simulate from PKPD model to get estimates of target attainment.

- Assume a relationship between microbiological kill and clinical endpoint based upon literature review.

- Based upon this relationship, provide estimate of cure probability & its uncertainty
  - Construct prior distribution to represent this estimate.
  - Dependant on confidence in translation from micro kill to clinical endpoint, different levels of uncertainty introduced into prior distribution

- Risks
  - No relationship between microbiological cure and clinical endpoint
  - PKPD model does not apply in this situation



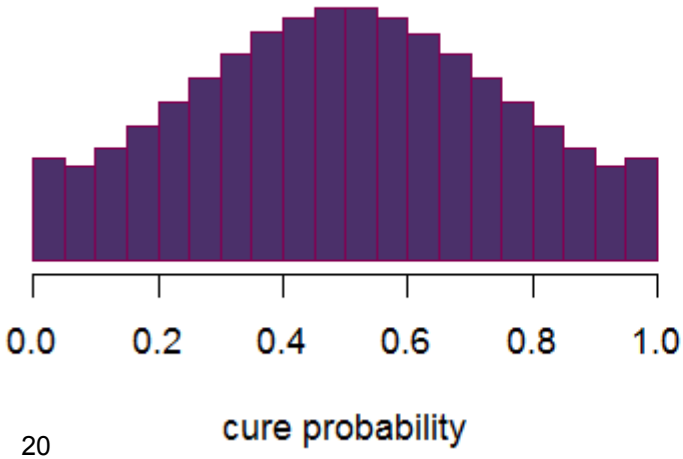
# Bayesian Approaches

## Use of PK/PD data to construct prior distributions

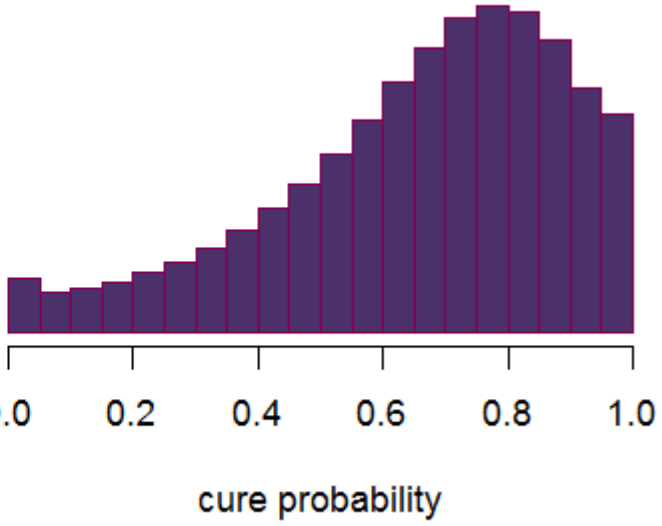
Worked Example

- 240 patients randomised overall
  - 2:1 ratio; 160 experimental v 80 control)
  - 24 patients on experimental and 12 on control with known positive MDR status
- Cure probabilities
  - 78% experimental v 76% control for non-MDR pathogens
  - 66% v 64% for MDR pathogens
- Prior distribution applied to both treatment arms

Broad Prior – centred around 50% response




Stronger Prior – centred around 75% response



# Bayesian Approaches

## Use of PK/PD data to construct prior distributions

### Results

**80% confidence/predictive intervals for cure probabilities**

| | Method | 80% interval for cure probabilities | | 80% interval for difference in cure probability |
|---|---|---|---|---|
| | | Experimental | Control | |
| Overall population (n=160 exp vs. 80 cont) | Frequentist | (0.67, 0.77) | (0.62, 0.77) | (-0.06, 0.11) |
| | Bayes (broad prior) | (0.66, 0.75) | (0.61, 0.74) | (-0.05, 0.11) |
| | Bayes (stronger prior) | (0.68, 0.76) | (0.65, 0.76) | (-0.05, 0.09) |
| MDR-positive patients (n=24 exp vs. 12 cont) | Frequentist | (0.47, 0.76) | (0.33, 0.67) | (-0.16, 0.41) |
| | Bayes (broad prior) | (0.52, 0.76) | (0.40, 0.68) | (-0.11, 0.23) |
| | Bayes (stronger prior) | (0.57, 0.79) | (0.48, 0.75) | (-0.13, 0.20) |

*Exp = experimental; Cont = Control*



# Bayesian methods

*Bayesian approaches provide useful techniques, but it is critical to understand assumptions underpinning their use*

Points to consider:

- Goodness of fit from any Bayesian predictions or models
- Strength of prior & how influential this is *vs.* observed data
- External data prior: need confidence in similarity of design, patient population, anticipated effects
- PK/PD prior: concentration levels not randomly assigned
  - Patients with different concentrations may differ on other factors which impact response (age, severity, co-morbidities)
- For rare pathogens, approaches may help quantify available data, but the influence of the prior distribution should be considered carefully



# Ideas in this talk

- What is the issue and how could we approach it?

- What are the options when only smaller RCTs are possible?
  - Statistical criteria
  - Bayesian approaches

- Interpretation of information on small numbers of resistant pathogens
  - Data presentation when inferential testing is challenging
  - Use of supplemental information from external sources
  - Issues and methods with using all available information



# When only limited data are possible…

- In a situation where a small microbiologically confirmed population is of primary interest

- There seem to be two options
  - A *very* small RCT (so small that inferential testing is not possible)
  - *Open-label* data (single arm trial)

  *For both approaches, external data could be used to set minimal efficacy levels*

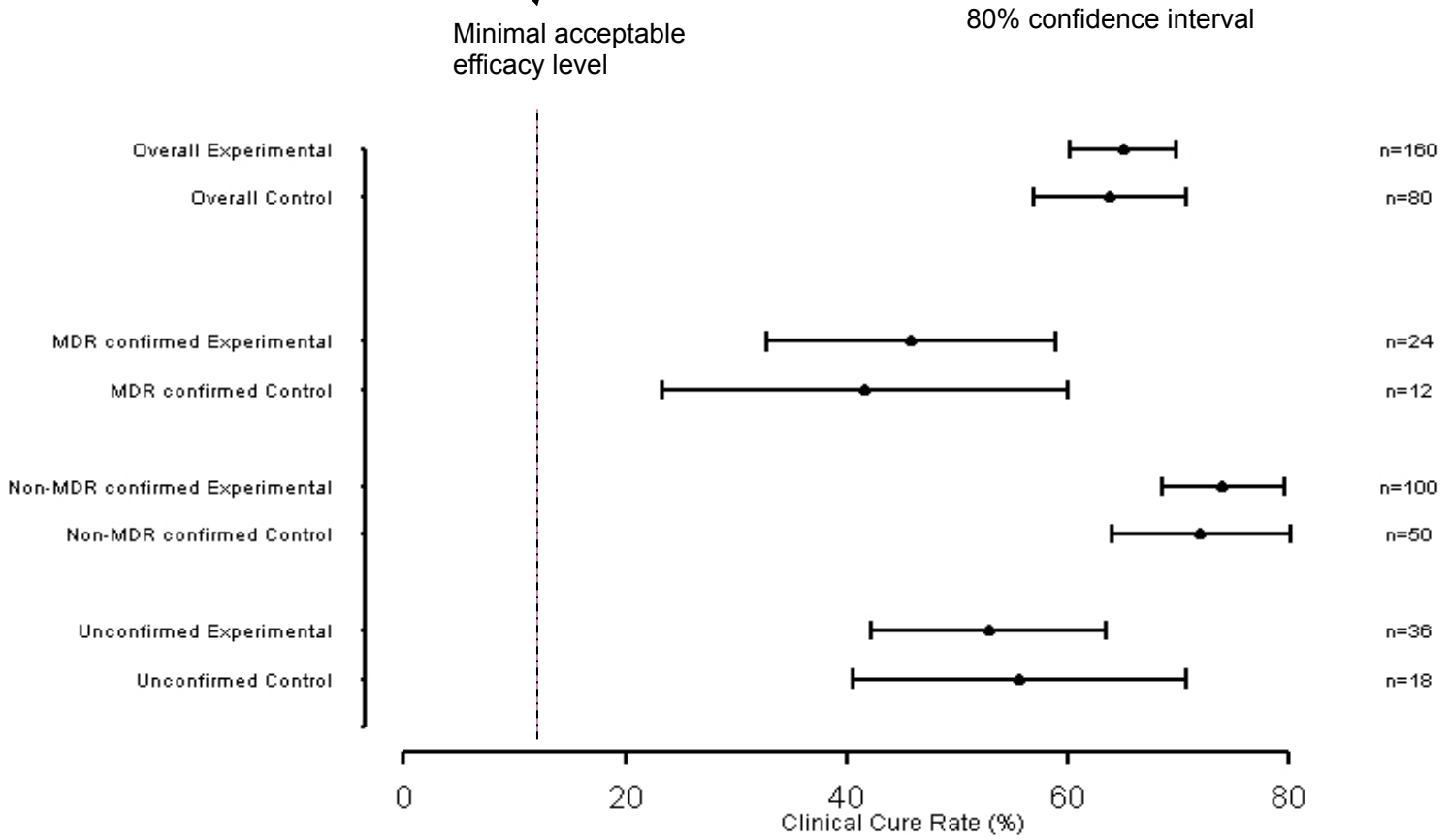
- Points to consider on single arm data
  - Small RCT gives randomisation, but heterogeneity may lead to problems of comparability of treatment groups
  - Non randomised study leads to concerns of comparability with externally generated data
  - Optimal route depends on quality & nature of external data available



# Data Presentation From Smaller RCT datasets

**Use Comparator data to provide context for the disease setting in question**

**Where possible, include a reference to a minimal level of efficacy based on a clinical justification and/or external data to give confidence of actvity**



# External Controls – key issues to consider

- Key question: Are data available in appropriate population?

- Contemporary controls most useful because resistance patterns, supportive care and other factors are changing

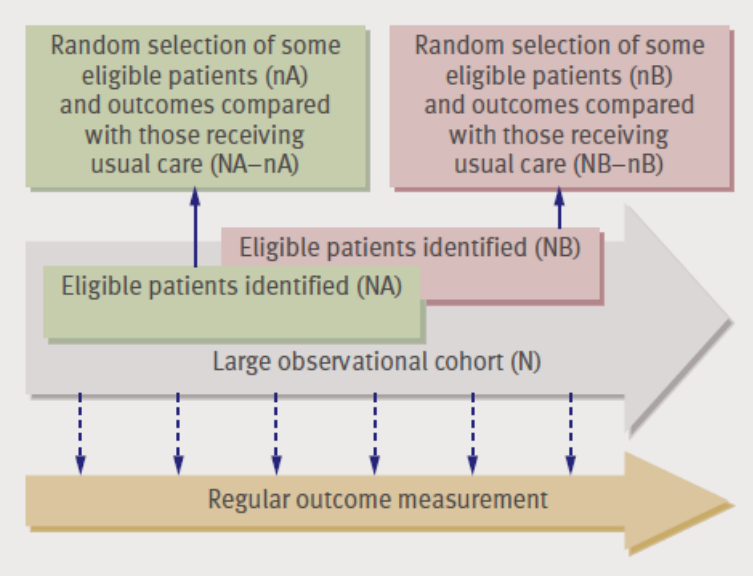
- But: does it really make development more feasible?
  - Historical controls may improve feasibility, but are available data appropriate?
  - Prospective data allow designs similar to RCTs, but face similar issues of patient availability as RCT
  - Prospective data generation on SOC during earlier phases of development may help

***- External data should be considered, but needs to be feasible and relevant***

***- Further discussion on possible sources of external data may help***

# An Alternative Approach Using Observational Data

## A Design That Warrants Further Consideration




The "cohort multiple randomised controlled trial" design. Firstly, a large observational cohort of patients with the condition of interest is recruited (N) and their outcomes regularly measured. Then for each randomised controlled trial, information from the cohort is used to identify all eligible patients (NA). Some eligible patients (nA) are randomly selected and offered the trial intervention. The outcomes of these randomly selected patients (nA) are then compared with the outcomes of eligible patients not randomly selected; that is, those receiving usual care (NA – nA). This process can be repeated for further randomised controlled trials (for example, NB)

**RESEARCH METHODS & REPORTING**

Rethinking pragmatic randomised controlled trials: introducing the "cohort multiple randomised controlled trial" design

Clare Relton,[13] David Torgerson,[2] Alicia O'Cathain,[1] Jon Nicholl[1]

BMJ | 1 MAY 2010 | VOLUME 340

**The Idea**

- **Arrange for routine collection of specific MDR pathogens**
- **Randomly select a cohort of eligible patients for test treatment(s)**
- **A number of treatments in development could utilise a dataset in this way**

**To consider:**

- **Is prospective identification of patients with MDR pathogen feasible in terms of timing from pathogen identification to inclusion in this trial?**
- **Could these ideas be applied in a different way to MDR pathogen trials?**

# Conclusion

- Traditional statistical inference not possible in some settings
  - Agree nature of unmet need with agency at outset and define approach accordingly
  - Further discussion and evaluation of alternative techniques will help
- Consider use of external data sources, but with care
- Balance of uncertainty & feasibility for areas of unmet need
  - But changes in uncertainty should be distinguished from areas which could potentially bias interpretation.

***Alternative strategies are critical to ensure a path forward which is both feasible and acceptable to regulatory agencies***



## Confidentiality Notice