CTTI Statistics Think Tank for Anti-Bacterial Drug Development

Lisa LaVange

Office of Biostatistics, CDER, FDA

August 20, 2012 Bethesda Hyatt

Background

- Recent advances in clinical trial design in the past decade have offered new and accelerated pathways for drug development in some therapeutic areas
  - Examples include adaptive designs, targeted enrollment based on genomic and proteomic testing, etc.
- During the same period, advances have also been made in understanding the natural history of acute bacterial infections
- In spite of these advancements, clinical trial design and conduct for anti-bacterial therapies remains challenging

Background

- Economic challenges include treatment periods of short duration and narrowly defined patient populations
- The scientific and economic challenges have combined to reduce the level of anti-bacterial drug development activity and shrink the pipeline of promising new therapies
- As resistance to current antibiotics increases, the absence of new therapies becomes critical

Objectives

- FDA is undertaking a number of initiatives to promote anti-bacterial drug development
- CTTI Statistics Think Tank provides an opportunity for leading experts in clinical trial methodologies to discuss alternative approaches to design and analysis that may prove useful for anti-bacterial programs
- Goal is to increase the likelihood that clinical trials of promising agents are successful and to ensure that those agents, if approved, are in fact safe and effective therapies for the intended patient populations

Specific Issues

- Enrollment
  - There is often limited time to recruit, enroll, and administer treatment due to severity of infection
  - Enrollment may take place in emergency rooms, an often difficult recruitment setting
  - Patients may need immediate treatment, prior to randomization, and the impact of that prior therapy in the ability to assess efficacy of the test agent can be considerable

  ➔ Efficient designs are needed to minimize sample size required

Specific Issues

- Non-inferiority designs
  - In many cases, it is unethical to use placebo as control, resulting in the need for non-inferiority trial designs
  - At the same time, there may be limited historical data on placebo response rates, making the identification and justification of a non-inferiority margin difficult
  - Other well-known issues with non-inferiority trials come into play, perhaps to a greater degree than in other disease areas, e.g., the tendency for other sources of variability to mask important differences between test treatment and control

Favorable Aspects

- Availability of prior information
  - Pre-clinical data may be able to confirm that the pharmacological agent is effective against the pathogen in vitro
  - PK/PD data can provide important information about availability of the agent in vivo, leading to optimal dosing
  - Both sources combine to facilitate phase 2/3 study designs
- Availability of information from other sources
  - The agent may have already been studied or even approved for use against the same pathogen but at a different (body) site of infection
  - Similarly, the agent may have known characteristics in combatting other pathogens at the same (body) site

Favorable Aspects

- Availability of information from other sources, cont.
  - Even though historical data on placebo-treated patients may be lacking, there may be lots of data from clinical trials in the same indication and similar patient populations on active control agents

➔Efficient ways to take advantage of prior or external data are needed

Plan for the Day

- FDA statistics review team identified four broad areas to focus the discussion
  1. **One- versus two-study paradigm:** When does it make sense to plan for a single, confirmatory study as sufficient evidence of efficacy and safety in treating antibacterial infections, and what particular requirements should be placed on such a study, when planned?
  2. **Non-inferiority trials**: Are there more efficient ways to establish non-inferiority to an existing therapy, when placebo controlled studies are not ethical/possible, given the many challenges in this disease area?

Plan for the Day

- Discussion areas, cont.
  3. **Multiple (body) sites of infection**: are there efficient ways to combine information across multiple (body) sites for a single pathogen, or across multiple pathogens for a single (body) site to better inform confirmatory trial designs and analysis?
  4. Time permitting, are there innovative ways to approach a variety of other problems with anti-bacterial trials, e.g., accounting for prior therapies that cannot be withheld, handling missing data, dealing with the use of subpopulations as primary (due to delays in confirming pathogen causing infection), etc.?

Plan for the Day

- The day is divided into four discussion sessions corresponding to each of these broad areas
- We will begin each session with a very brief (5-minute) background presentation to set the stage and pose specific questions
- At anytime during the discussion, please feel free to offer new proposals, react to proposals on the table, or contribute in any way you feel is productive
- We will try to stay to the schedule, in order for some discussion of each broad area

Success!

- A successful outcome for the day will
  - Be a lively exchange of ideas from varied perspectives (academic, industry, and regulatory)
  - Generate proposals for innovative study design and analysis that FDA statisticians can pursue for regulatory feasibility
  - Prompt continued discussion among participants post-meeting on research ideas of mutual interest
- Thank you for your willingness to participate and engage!

Session 1

One- vs two-study paradigm for approval

Lisa LaVange

Office of Biostatistics, CDER, FDA

Regulatory Standard For Effectiveness

- Substantial evidence – defined by the U.S. Food, Drug, and Cosmetic Act
  - "…evidence consisting of adequate and well-controlled investigations, including clinical investigations, …to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports…"
- FDA's interpretation of the statute
  - At least two "adequate and well-controlled" trials, each convincing on its own, are required to establish effectiveness.

1997 Amendments

- Food and Drug Amendments Act (FDAMA) states that FDA may consider
  - Data from one adequate and well-controlled clinical investigation and confirmatory evidence to constitute substantial evidence
  - If FDA determines that such data and evidence are sufficient to establish effectiveness

FDA Guidance Following FDAMA

- FDA Guidance for Industry (1998): Providing Clinical Evidence of Effectiveness for Human Drugs and Biologic Products
  - Describes circumstances in which FDA may rely on a single trial to demonstrate effectiveness for human drugs and biologic products
  - Currently under revision

Clinical Evidence from a Single Study

- Characteristics of a single, adequate, well-controlled trial to support an effectiveness claim
  - A large, multi-center trial in which no single site provided an unusually large fraction of the patients and no single investigator or site was disproportionally responsible for the favorable effect seen
  - Consistency of study findings across key patient subsets (e.g., disease stage, age, gender, race)
  - Presence of multiple studies within a single study, such as occurs in a factorial design, which show consistent findings
  - Persuasive evidence on multiple endpoints
  - A statistically very persuasive finding (evidence comparable to that found from two studies)

Clinical Evidence from a Single Study

- References
  - Shun, Chi, Durrleman, and Fisher (2005), Statistical consideration of the strategy for demonstrating clinical evidence of effectiveness – one larger vs two smaller pivotal studies, Statistics in Medicine
  - Commentary on the above by Gary Koch (2005), Statistics in Medicine
  - Commentary on the above by Mohammad Huque (2005), Statistics in Medicine

Session 2a

Non-inferiority (NI) trial designs, choice of margins, and analysis strategy: Nosocomial Pneumonia

Scott Komo, DrPH

Office of Biostatistics, CDER, FDA

Outline

- Literature search
- Historical evidence
- Selection of studies
- Estimation of active control treatment effect
- NI margin determination
- Extrapolation of the treatment effect in all-cause mortality to clinical response

Historical Evidence

- Original journal articles (1970-2008)
- No placebo-controlled clinical trials
- Placebo effect for all-cause mortality estimated indirectly:
  - 12 studies of patients administered inappropriate, delayed, or inadequate initial treatment that reported all-cause mortality
    - Non-randomized, observational cohort studies
- Active control effect:
  - 9 randomized, active-controlled clinical trials
    - Primary endpoint: Clinical response
    - Secondary endpoint: all-cause mortality

Selection of Studies

- Comparability of groups

  Selected a subset of studies due to concerns on the comparability of patients based on

  – Age

  – Severity of Illness

- Placebo

  Selected 2 out of 12 studies

- Active control

  Selected 5 out of 9 studies

Estimation of the Active Control Treatment Effect

- Fixed margin approach
- Estimated the placebo and active control mortality rates separately using DerSimonian and Laird random effects meta-analyses
  - Placebo mortality rate: 62%; (52%,71%)
  - Active control mortality rate: 20%; (18%, 23%)
- Active control treatment effect estimate: 29%
  [52% - 23%]

NI Margin Determination

- A 10% NI margin was felt to be justifiable given the large active control treatment effect

- There are concerns using an NI margin of greater than 10% for a mortality endpoint

Can we extrapolate the treatment effect seen in all-cause mortality to justify an NI margin for clinical response?

Mortality vs. Clinical Response

- All-cause mortality
  - Clinically critical
  - Placebo effect could be estimated from patients who received inappropriate/delayed/inadequate initial treatment
  - Concerns:
    - Noise due to non-infection related deaths
    - Window to capture deaths is not entirely clear
    - Possible effect of the discontinuation of life support
- Clinical response
  - Most clinicians prefer this endpoint to assess efficacy
  - No historical placebo data
  - Definition of Failure (At the End of Therapy (EOT) or a predefined period after EOT
    - Lack of resolution of clinical signs and symptoms of pneumonia OR
    - Died
  - Active control trials have data for both clinical response and mortality

Questions

- **Question 1:** What methods or what types of data are needed to be able to translate or bridge margins from one endpoint (e.g. mortality) to another (e.g. clinical response)? Would case-control studies, for example, provide the additional information needed? Can the estimation of correlation between endpoints from other studies be helpful in this regard?

- **Question 2**: What are the advantages/disadvantages of other approaches to margin determination in regulatory studies, e.g., a Bayesian approach?

- **Question 3:** Are there efficiencies to be gained through the use of other analysis methods, such as Bayesian analysis (e.g. Gamalo, Wu, and Tiwari), and if so, at what cost?

Backup Slides

Inadequate/Delayed therapy: All-Cause Mortality Rate

Studies	# with NP	Mean Age SD, Years	Severity (Mean APACHE II SD)	All-cause Mortality n/N (%)
[1]	130	Inadequate 53.0 17.7	17.5 4.9	31/51 (61%)
[2]	76	Delayed 66 17	19 6	33/52 (64%)

Active Control: All-Cause Mortality Rate

Studies	ITT N	Active Control groups	Mean Age SD, Years	Mean APACHE II score SD	All-cause Mortality (ITT), n/N (%)
[3]	124	P/T/A Cef/A	P/T/A: 57.1 17 Cef/A: 60.5 20	P/T/A: 16.5 6.6 Cef/A: 16.9 6.5	P/T/A: 27/88 (31%) Cef/A: 8/36 (22%)
[4]	402	Cip Imi	Cip: 59.9 17.9 Imi: 59.6 17.6	Cip: 17.7 6.5 Imi: 17.6 6.4	Imi: 38/200 (19%) Cip: 43/202 (21%)
[5]	438	Lev iv /Lev po Imi iv/Cip po	Lev iv/Lev po: 55.8 20.0 Imi iv/Cip po: 55.5 20.1	Lev iv/Lev po: 15.0 5.8 Imi iv/Cip po: 14.8 6.0	Lev iv/Lev po: 38/220 (17%) Imi iv/Cip po: 32/218 (15%)
[6]	396	LZD/AZM Van/AZM	LZD/AZM: 62.8 18.0 Van/AZM: 61.3 18.7	LZD/AZM: 15.7 6.5 Van/AZM: 15.4 6.9	LZD/AZM: 36/203 (18%) Van/AZM: 49/193 (25%)
[7]	623	LZD/AZM Van/AZM	LZD/AZM: 63.1 19.1 Van/AZM: 61.9 19.3	LZD/AZM: 14.1 5.8 Van/AZM: 14.1 6.2	LZD/AZM: 64/321 (20%) Van/AZM: 61/302 (20%)

P/T=piperacillin/tazobactam; A=amikacin; Cef=ceftazidime; TO=tobramycin; Cip=ciprofloxacin; Imi=imipenem; Lev=levofloxacin; LZD=linezolid; Van=vancomycin; AZM=aztreonam; SD=standard deviation

Session 2b

Bayesian Approach to Meta-analysis and Non-inferiority Trials

M. Amper Gamalo, PhD
and
Ram C. Tiwari, PhD
Office of Biostatistics, CDER, FDA

Random Effects Meta-analysis

- Assume that there are k studies, and that

$$Y_i = \mu + \alpha_i + \epsilon_i, \tag{1}$$

where $\epsilon_i \overset{ind}{\sim} N(0, \sigma_i^2)$

- $\sigma_i^2 > 0,\ i = 1, \ldots, k$ (known) are the within study variabilities
- α_i are the random effects $\overset{iid}{\sim} N(0, \tau^2)$
- $\tau^2 > 0$ (unknown) is the between study variability
- $\tau^2 = 0$ implies that studies are homogeneous, and $\tau^2 > 0$ implies that studies are heterogenous

- Full model

$$\begin{aligned} Y_i|\mu, \alpha_i &\overset{ind}{\sim} N(\mu + \alpha_i, \sigma_i^2),\ \sigma_i^2 > 0 : \text{known}; \\ \alpha_1, \ldots, \alpha_k|\tau^2 &\overset{iid}{\sim} N(0, \tau^2),\ \tau^2 \geq 0 : \text{unknown} \end{aligned} \tag{2}$$

Frequentist Estimate

- Estimate for μ

$$\hat{\mu} = \frac{\sum_i \hat{w}_i^* Y_i}{\sum_i \hat{w}_i^*}; \quad s.e.(\hat{\mu}) = \left\{ \sum_i \hat{w}_i^* \right\}^{-\frac{1}{2}} \tag{3}$$

- The weights $\hat{w}_i^*$ are obtained from $w_i^* = 1/(\sigma_i^2 + \tau^2)$ and the DerSimonian-Laird estimate of τ^2

$$\hat{\tau}_{DL}^2 = \max\left(0, \frac{Q - (k-1)}{\sum_i w_i - (\sum_i w_i^2 / \sum_i w_i)}\right) \tag{4}$$

where $Q = \sum w_i (Y_i - \hat{\mu}_{MH})^2$, $\hat{\mu}_{MH} = \sum_i w_i Y_i / \sum_i w_i$, and $w_i = \sigma_i^{-2}$

Bayesian Estimation using Normal Prior

- Random effects meta-analysis with Normal Prior

$$Y_i|\mu,\alpha_i \overset{ind}{\sim} N(\mu+\alpha_i,\sigma_i^2),\sigma_i^2 : \text{known} \tag{5}$$

Priors:

$$\mu \sim \text{improper}, \quad \alpha_i|\tau^2 \overset{iid}{\sim} N(0,\xi^{-1}\tau^2), \quad i=1,\ldots,k, \quad \xi : \text{(un?)known} = 1$$

$$\mu \coprod \alpha_i|\tau^2, \qquad \tau^2 \sim IG$$

- The normality assumption on α may be too strong when there is considerable heterogeneity among studies

Bayesian Estimation Using Dirichlet Process Prior

- Sethuraman and Tiwari (1981) and Sethuraman (1984):

$$G = \sum_{k=1}^{\infty} \pi_k \delta_{\alpha_k}; \quad \alpha_k \overset{iid}{\sim} H \tag{7}$$

$$\pi_1 = \theta_1; \quad \pi_k = \theta_k \prod_{j=1}^{k-1} (1-\theta_j); \quad k \geq 2 \tag{8}$$

$$\theta_k \overset{iid}{\sim} \text{Beta}(1, \rho) \tag{9}$$

- Finite representation (easy to implement in Winbugs)

$$G = \sum_{k=1}^{L} \pi_k \delta_{\alpha_k}; \quad \pi_L = 1 - \pi_1 - \ldots - \pi_{L-1} \tag{10}$$

The Stick-Breaking Representation of Dirichlet Process

- Sethuraman and Tiwari (1981) and Sethuraman (1984):

$$G = \sum_{k=1}^{\infty} \pi_k \delta_{\alpha_k}; \quad \alpha_k \overset{iid}{\sim} H \tag{7}$$

$$\pi_1 = \theta_1; \quad \pi_k = \theta_k \prod_{j=1}^{k-1} (1-\theta_j); \quad k \geq 2 \tag{8}$$

$$\theta_k \overset{iid}{\sim} \text{Beta}(1, \rho) \tag{9}$$

- Finite representation (easy to implement in Winbugs)

$$G = \sum_{k=1}^{L} \pi_k \delta_{\alpha_k}; \quad \pi_L = 1 - \pi_1 - \ldots - \pi_{L-1} \tag{10}$$

HAP/VAP Data

Inadequate or delayed therapy all-cause mortality rates

Studies	ITT (N)	Therapy	All-cause Mortality, n/N (%)
[1]	130	Inadequate	31/51 (61%)
[2]	76	Delayed	33/52 (64%)

Active-control all-cause mortality rates

Studies	ITT (N)	Active Control groups	All-cause Mortality, n/N (%)
[3]	124	P/T/A	27/88 (31%)
		Cef/A	8/36 (22%)
[4]	402	Cip	43/202 (21%)
		Imi	38/200 (19%)
[5]	438	Lev iv/Lev po	38/220 (17%)
		Imi iv/Cip po	32/218 (15%)
[6]	396	LZD/AZM	36/203 (18%)
		Van/AZM	49/193 (25%)
[7]	623	LZD/AZM	64/321 (20%)
		Van/AZM	61/302 (20%)

Meta-analysis of HAP/VAP Data

	Frequentist: DerSimonian-Laird % (95% Conf. Interval)	**Bayesian: Dirichlet Process Prior % (95% Cred. Interval)**
Inadequate or Delayed Therapy	0.62 (0.52, 0.71)	0.62 (0.53, 0.72)
Active Control	0.20 (0.18, 0.23)	0.20 (0.18, 0.21)

- Treatment effect through frequentist (DerSimonian-Laird) method: $52\% - 23\% = 29\%$
- Treatment effect through Bayesian (Dirichlet Process Prior) method: $53\% - 21\% = 32\%$
- Calculation of the Treatment effect is based on the recommendation given in the Non-inferiority Guidance
- Data suggests that 10% margin is justified

Frequentist Decision Rule

- Hypothesis

$$H_0 : \mu_E - \mu_C \leq -\delta \quad vs. \quad H_0 : \mu_E - \mu_C > -\delta \tag{11}$$

where μ_E is the mean effect of the experimental drug in the current trial; μ_C is the mean effect of the control drug in the current trial; δ is the pre-specified margin

- Let $X_{E,i}$, $X_{C,j}$, $(i = 1, 2, ..., n_E; j = 1, 2, ..., n_C)$ denote the random variables corresponding to the experimental and reference treatment responses in the current non-inferiority trial, respectively. Then, the frequentist decision rule is to reject the null if

$$(\bar{X}_E - \bar{X}_C) - z_{1-\alpha/2}\sqrt{\frac{\sigma_E^2}{n_E} + \frac{\sigma_C^2}{n_C}} > -\delta \tag{12}$$

- In the presence of unequal variance, fixed level tests are not available. Test can be based on Welch test or generalized p-value approach (Gamalo and Tiwari, 2011).

Bayesian Approach

- Assume that μ_E has a non-informative prior given by $\pi(\mu_E) \propto 1$
- Let the prior for μ_C be informative given by $\mu_C \sim N\left(\mu_C^* = \bar{X}_{C_0}, \sigma_C^{*2} = \sigma_{C_0}^2/n_{C_0}\right)$
- The posterior distributions of μ_E and μ_C are

$$\mu_E|\bar{X}_E, \sigma_E^2 \sim N\left(\bar{X}_E, \frac{\sigma_E^2}{n_E}\right) \tag{13}$$

$$\mu_C|\bar{X}_C, \sigma_C^2 \sim N(\tilde{\mu}_C, \tilde{\sigma}_C^2) \tag{14}$$

- Estimates for $\tilde{\mu}_C$ and $\tilde{\sigma}_C^2$ are

$$\tilde{\mu}_C = \tilde{\sigma}_C^2\left(\frac{n_c\bar{x}_C}{\sigma_C^2} + \frac{\mu_C^*}{\sigma_C^{2*}}\right) = \tilde{\sigma}_C^2\left(\frac{n_c\bar{x}_C}{\sigma_C^2} + \frac{n_{C_0}\bar{x}_{C_0}}{\sigma_{C_0}^2}\right) \tag{15}$$

$$\tilde{\sigma}_C^2 = \left(\frac{n_C}{\sigma_C^2} + \frac{1}{\sigma_C^{2*}}\right)^{-1} = \left(\frac{n_C}{\sigma_C^2} + \frac{n_{C_0}}{\sigma_{C_0}^2}\right)^{-1} \tag{16}$$

Bayesian Decision Rule

- Reject H_0 if:

$$p(\mu_E - \mu_C \geq -\delta | \bar{X}_E, \bar{X}_C, \sigma_E^2, \sigma_C^2) = P\left(Z \geq \frac{-\delta - (\bar{X}_E - \tilde{\mu}_C)}{\left(\frac{\sigma_E^2}{n_E} + \tilde{\sigma}_C^2\right)^{\frac{1}{2}}} \right) \geq p^* \quad (17)$$

- Equivalently, reject H_0 if:

$$\bar{X}_E - \tilde{\mu}_C - z_{1-\alpha/2}\sqrt{\frac{\sigma_E^2}{n_E} + \tilde{\sigma}_C^2} > -\delta \quad (18)$$

for p^*. (Gamalo et al., 2012)

Application of Bayesian Decision Rule: Example 1

Ceftaroline vs. Ceftriaxone for CABP				
	Study 08		Study 09	
	Ceftaroline	Ceftriaxone	Ceftaroline	Ceftriaxone
All-Cause Mortality	4/291	5/300	7/284	5/269
Observed Proportion	1.4%	1.7%	2.5%	1.9%
Diff (95% Conf Int)	-0.3% (-2.9, 2.3)		0.6% (-2.4, 3.6)	
Posterior Mean	1.4%	2.6%	2.5%	3.0%
Diff (95% Cred Int)	-1.3% (-3.1, 0.6)		-0.5% (-2.8, 1.8)	

- The prior used for ceftriaxone was based on the all-cause mortality rate of 7.8% ($n = 243$) obtained from CABP Guidance Table 4.

Application of Bayesian Decision Rule: Example 2

	Study 001		Study 002	
	Experimental	**Active Control**	**Experimental**	**Active Control**
All-Cause Mortality	90/400	70/390	65/350	75/370
Observed Proportion	22.5%	17.9%	18.6%	20.3%
Diff (95% Conf Int)	4.6% (-1.2, 10.3)		-1.7% (-7.7, 4.3)	
Posterior Mean	22.5%	21.2%	18.6%	21.7%
Diff (95% Cred Int)	1.3% (-2.9, 5.4)		-3.1% (-8.0, 1.7)	

- Historical active control all-cause mortality rate of 22%, s.d. = 0.01862.

Question for Discussion

- **Question 1:** What are the advantages/disadvantages of other approaches to margin determination in regulatory studies, e.g. a Bayesian approach?

- **Question 2:** Are there efficiencies to be gained through the use of other analysis methods, such as Bayesian analysis (e.g. Gamalo, Wu, Tiwari), and if so, at what cost?

- **Question 3:** How does one make a decision when faced with differing results from the Bayesian and frequentist methods?

References

- Gamalo, M. and Tiwari, R. (2011) *Testing Noninferiority Through Generalized p-values.* Statistics in Biopharmaceutical Research **3***: 87-96.*
- Gamalo, M., Wu , R., Tiwari, R. (2011) *Bayesian Approaches to noninferiority testing with binary endpoints.* Journal of Biopharmaceutical Statistics **21**: 902-919
- Gamalo, M., Wu, R., Tiwari, R. *Bayesian Approaches to noninferiority trials with normal means. Stat Methods Med Res* published online 21 May 2012
- Muthukumarana, S. and Tiwari, R. Meta-analysis using Dirichlet Process. *Stat Methods Med Res* published online 16 July 2012
- Sethuraman, J. and Tiwari, R. Convergence of Dirichlet measures and the interpretation of their parameter. *Stat Dec Theo and Related Topics III*, Academic Press, 1982
- Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Stat Sin*, 1994

Discordant MIC Analysis: A New Path for Licensure of Anti-infective Drugs

Dean Follmann,
Erica Brittain, and John Powers
NIAID

Current NI Trial Paradigm

- 1) Confidence interval for the difference in success rates for New Drug B – Comparator Drug A lies to the right of a margin **M.**

PLUS

- 2) **M** based on historical evidence of the magnitude of the benefit of A versus placebo, tempered with clinical judgment
  - No historical evidence, no path forward.

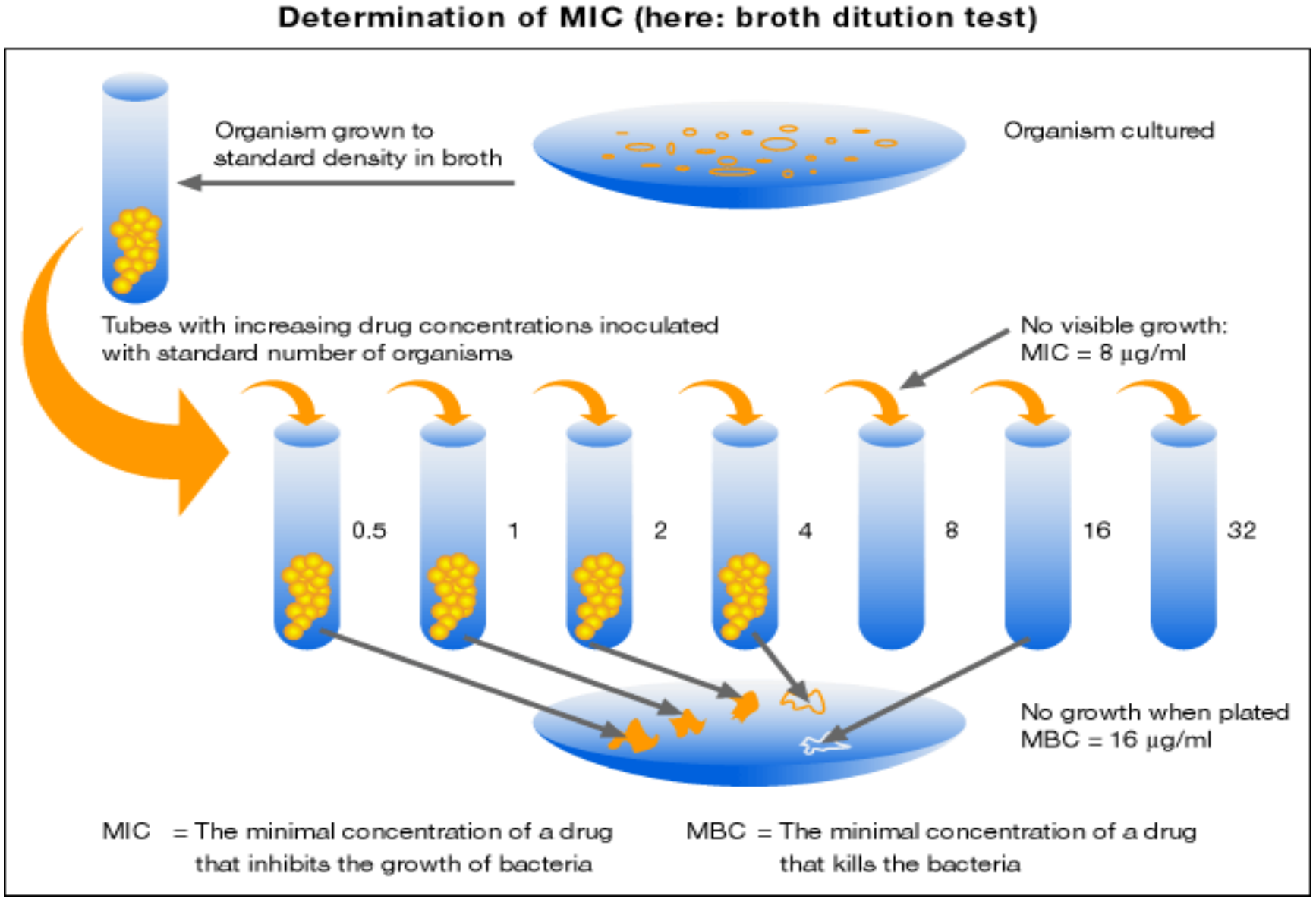
New NI Trial Paradigm

- 1) Confidence interval for the overall difference in success rates for New Drug B – Comparator Drug A lies to the right of a margin **M** based on clinical judgment.

PLUS

- 2) Superiority of B to A shown in pre-specified patients in current NI trial.
  - No need for historical evidence.

MIC measures Drug efficacy *in vitro*

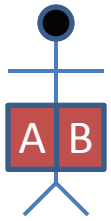


Drug A versus Drug B Clinical Trial
4 Kinds of People

Low MIC-A Low MIC-B

Low MIC-A High MIC-B

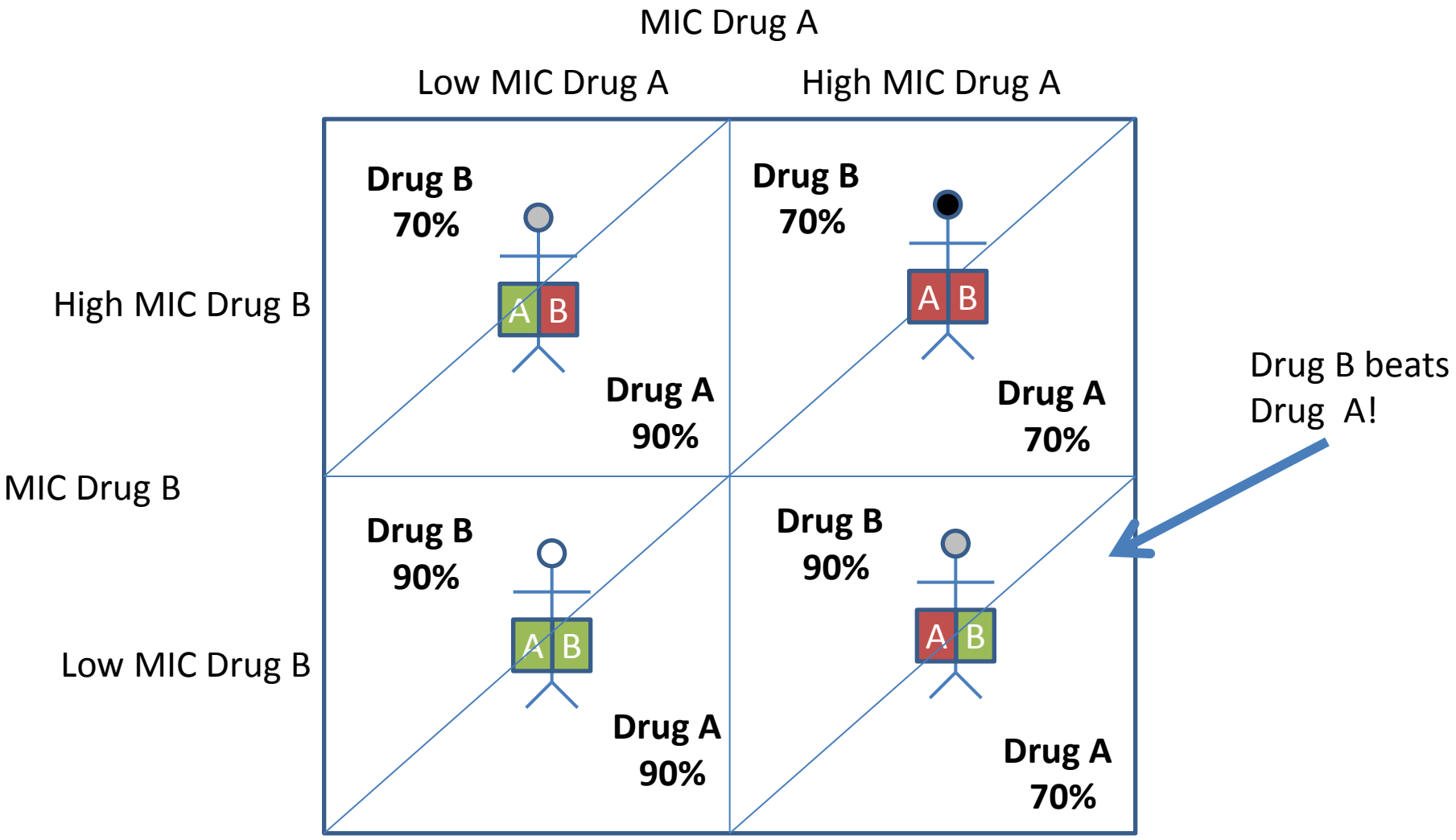
High MIC-A Low MIC-B**



High MIC-A High MIC-B

** Drug B should be superior to Drug A for these patients

The Key Subgroup Analysis



Discordant Regression Method

- Low B/High A patients may be rare.
- Use Logistic Regression to estimate the response surface.
- Log odds of success on B to success on A:

$$\beta_0 + \beta_1 Z + \beta_2 \text{MICA} + \beta_3 \text{MICB} + \beta_4 Z\,\text{MICA} + \beta_5 Z\,\text{MICB}$$

  - Z= 1 drug B (0 Drug A)

Test of Superiority of B over A

- *A priori*, find the (MIC-A, MIC-B) = (a,b) that maximally favors Drug B. See if B beats A at (a,b)---a fixed constant.

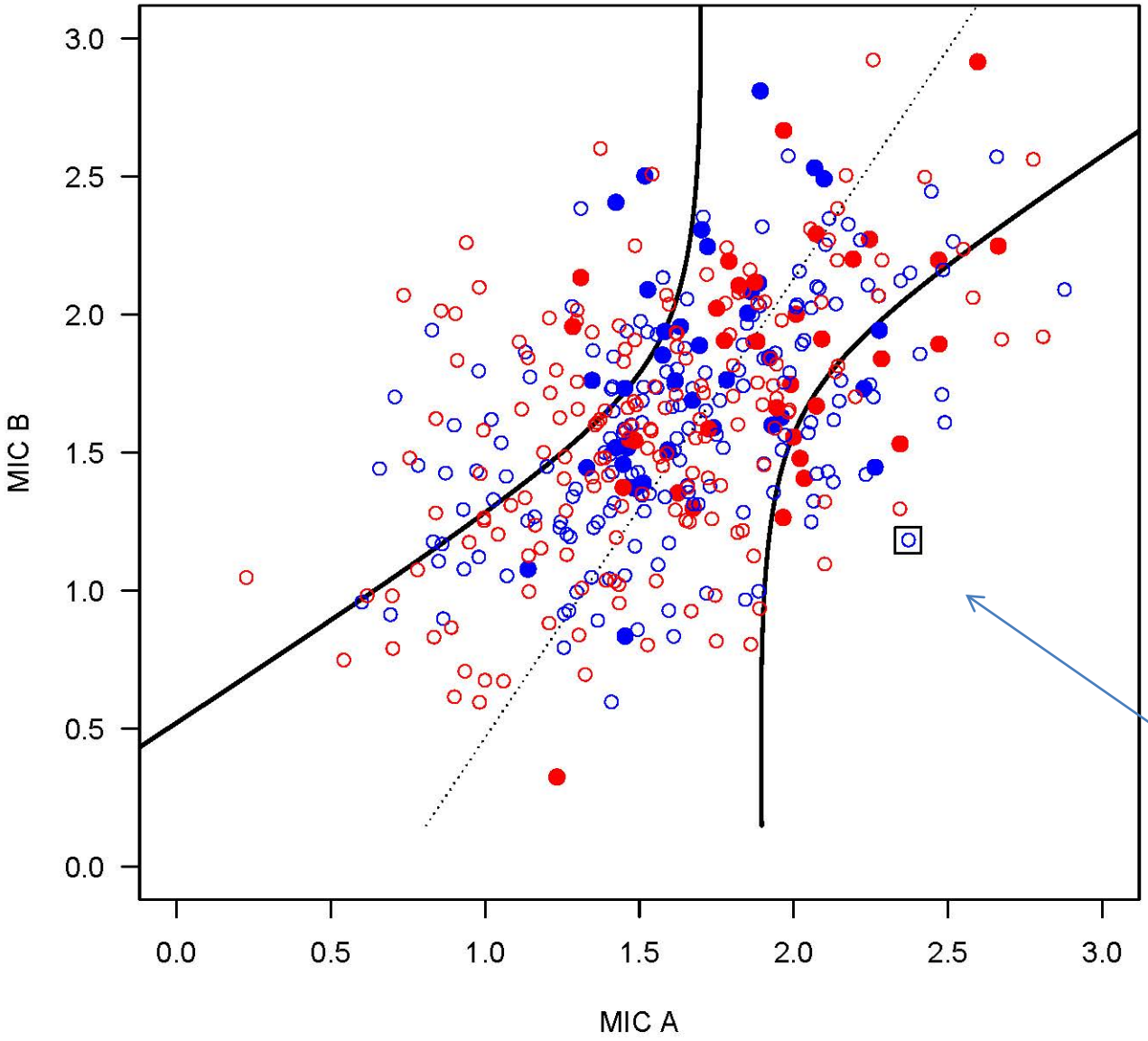
- Test H0: $\beta_1 + \beta_4 a + \beta_5 b = 0$

- Reject H0, conclude B is superior to A.

One Simulated Trial

- 1) Success rates .83, .82 for drug A, drug B 95% CI = (-.08,.07) NI for B with 10% margin
- 2) Discordant regression analysis. Confidence interval at the sweet spot CI = (.08, .60).
- Licensure supported.

One Simulated Trial



Sweet Spot
B beats A
p = .01

Patient Got Drug B and had Success

Patient got Drug A and Failed

Summary

- New paradigm for licensure
  - 1) Pick a clinically acceptable margin
  - 2) Test for superiority where it's most likely.
- Obviates need for historical evidence which may be shaky, nonexistent.
- Encourages a careful design to show superiority for *a priori* selected patients.
- Current work: Better tests, extend to AUC:MIC.

Session 3

Development plans that span multiple infection (body) sites

Daniel B. Rubin, PhD

Office of Biostatistics, CDER, FDA

Unmet Need and Resistant Pathogens

- Resistant pathogens are the key problem, but can be rare enough that it's challenging to directly study drugs with activity against them.
- Several recent proposals from EMA, industry, and physician groups for studying treatments for serious or life-threatening infections due to multidrug-resistant Gram-negative pathogens:

  Alemayehu et al. (2012). A paradigm shift in drug development for treatment of rare multidrug-resistant Gram-negative pathogens. *Clinical Infectious Diseases*.

- Ideas in literature proposals and actual submissions include:
  - Greater reliance on *in vitro*, animal, and PK/PD data
  - Trials with external/historical controls
  - Observational studies followed by postmarketing requirements
  - Active-controlled superiority trials, but pooling over body sites

Pooling Body Sites

- Suppose a trial combines bloodstream infections, urinary tract infections, and respiratory infections that are due to the same pathogen, such as *Pseudomonas aeruginosa*.
- This reverses the traditional paradigm of conducting large trials in each disease, which have few subjects with any specific pathogen.
- On approval, suppose that labeling will state the drug is indicated for treatment in the diseases studied due to the pathogen of interest when alternative therapies are not available or are not appropriate.
- How should this be done if there is little power for each disease?

Daptomycin

Daptomycin is approved for skin and other infections but it does not work in respiratory infections. Deactivation by pulmonary surfactant was only discovered in animal models after community-acquired pneumonia trials failed in humans.

CABP Studies 05+08	Clinical Response		
ITT Population	Ceftriaxone	Daptomycin	Difference
	326/431 (77.4)	293/413 (70.9)	6.5 (0.6, 12.4)

Source: Pertel et al. (2008). Effects of prior effective therapy on the efficacy of daptomycin and ceftriaxone for the treatment of community-acquired pneumonia. *Clinical Infectious Diseases.*

Tigecycline

- New Dehli metallo-beta lactamase 1 (NDM1) enzyme discovered in 2008 makes bacteria resistant to all antibacterial drugs except polymyxins (very toxic) and **tigecycline**.

- Superbug slowly spreading from India to the rest of the world.

Tigecycline: 2010 FDA Mortality Warning

Patients with outcome of death by infection type

Infection Type	Tygacil deaths/total patients (%)	Comparator Antibiotics deaths/total patients (%)	Risk Difference* (95% Confidence Interval)
cSSSI	12/834 (1.4%)	6/813 (0.7%)	0.7 (-0.3, 1.7)
cIAI	42/1382 (3.0%)	31/1393 (2.2%)	0.8 (-0.4, 2.0)
CAP	12/424 (2.8%)	11/422 (2.6%)	0.2 (-2.0, 2.4)
HAP	66/467 (14.1%)	57/467 (12.2%)	1.9 (-2.4, 6.3)
Non-VAP†	41/336 (12.2%)	42/345 (12.2%)	0.0 (-4.9, 4.9)
VAP†	25/131 (19.1%)	15/122 (12.3%)	6.8 (-2.1, 15.7)
RP	11/128 (8.6%)	2/43 (4.7%)	3.9 (-4.0, 11.9)
DFI	7/553 (1.3%)	3/508 (0.6%)	0.7 (-0.5, 1.8)
Overall Adjusted	150/3788 (4.0%)	110/3646 (3.0%)	0.6 (0.1, 1.2) **

Observe that:

- Meta-analysis shows statistically significant increased risk
- Risk difference is positive for every infection type
- (Standardized) difference largest for ventilator-associated pneumonia

source: http://www.fda.gov/Drugs/DrugSafety/ucm224370.htm

Doripenem

Mortality	Doripenem	Zosyn	P-value (Fisher's exact test)
During IV therapy	21/223 (0.09)	9/223 (0.04)	0.04
Due to pneumonia	9/221 (0.04)	1/221 (<0.01)	0.02

- FDA-approved for several indications such as abdominal infections
  - Not approved in 2008 for ventilator-associated pneumonia, partially from mortality signals in phase 3 program
- Approved by EMA with postmarketing requirement. Sponsor increased dose from 500 mg to 1 g from PK/PD and began new trial.

Source: http://www.fda.gov/ohrms/dockets/ac/08/briefing/2008-4364b1-01-FDA.pdf

Doripenem

Post market ventilator-associated pneumonia trial halted early from excess mortality and numerically worse clinical cure rates.

Analysis Population	Doripenem Group %	Imipenem Group %	Difference %	2-sided 95% CI %
Clinical Cure Rates				
MITT	45.6	56.8	-11.2	-26.3 to 3.8
ME	49.1	66.1	-17	-34.7 to 0.8
All Cause 28-day Mortality Rate (MITT)	21.5	14.8	6.7	-5.0 to 18.5

2012 FDA safety notice: http://www.fda.gov/Drugs/DrugSafety/ucm285883.htm

Points to Consider

- In unmet need or organism-based trials pooling body sites, could Bayesian or other methods ensure risks of future daptomycins, tigecyclines, or doripenems are understood at specific body sites?

- If not, then how should drugs be studied for treating infections due to resistant pathogens when it is considered infeasible to study the drug separately for each body site?

- Numbers needed to harm (for mortality) may be small if relatively ineffective drugs are used empirically for common life-threatening infections due to susceptible pathogens.

- Appropriate strategies for labeling and postmarketing requirements are not yet understood for new paradigms with pooling of body sites.

Questions

- **Question 1**: What types of study designs, including multiple testing strategies, should be considered for a single submission that includes clinical trials conducted in multiple infection sites, taking into account different background rates at different sites, etc.?

- **Question 2:** How should data from those trials be synthesized during analysis?

Backup Slide

How would you interpret hypothetical trial on mortality from resistant pathogens?

Disease	Mortality		
	Standard of Care	New Drug	Difference (95% CI)
Bloodstream Infections	15/30 (50.0)	5/30 (16.7)	33.3 (9.8, 53.7)
Intra-Abdominal Infections	7/15 (46.7)	3/15 (20.0)	26.7 (-7.2, 55.4)
Hospital-Acquired Pneumonia	7/15 (46.7)	10/15 (66.7)	-20.0 (-50.7, 15.1)
Pooled	29/60 (48.3)	18/60 (30.0)	18.3 (0.9, 34.8)

Session 4

Other Design and Analysis Considerations

Daniel B. Rubin, PhD

Office of Biostatistics, CDER, FDA

Focus on Community-Acquired Bacterial Pneumonia (CABP)

- Many issues related to feasibility have dealt less with the numerical value of the margin than other design features meant to ensure trials can differentiate effective and ineffective treatments.

- Selected issues:
  - Prior antibacterial therapy
  - Microbiological enrichment
  - Patient severity
  - Definition and timing of endpoint

CABP: Prior Antibacterial Therapy

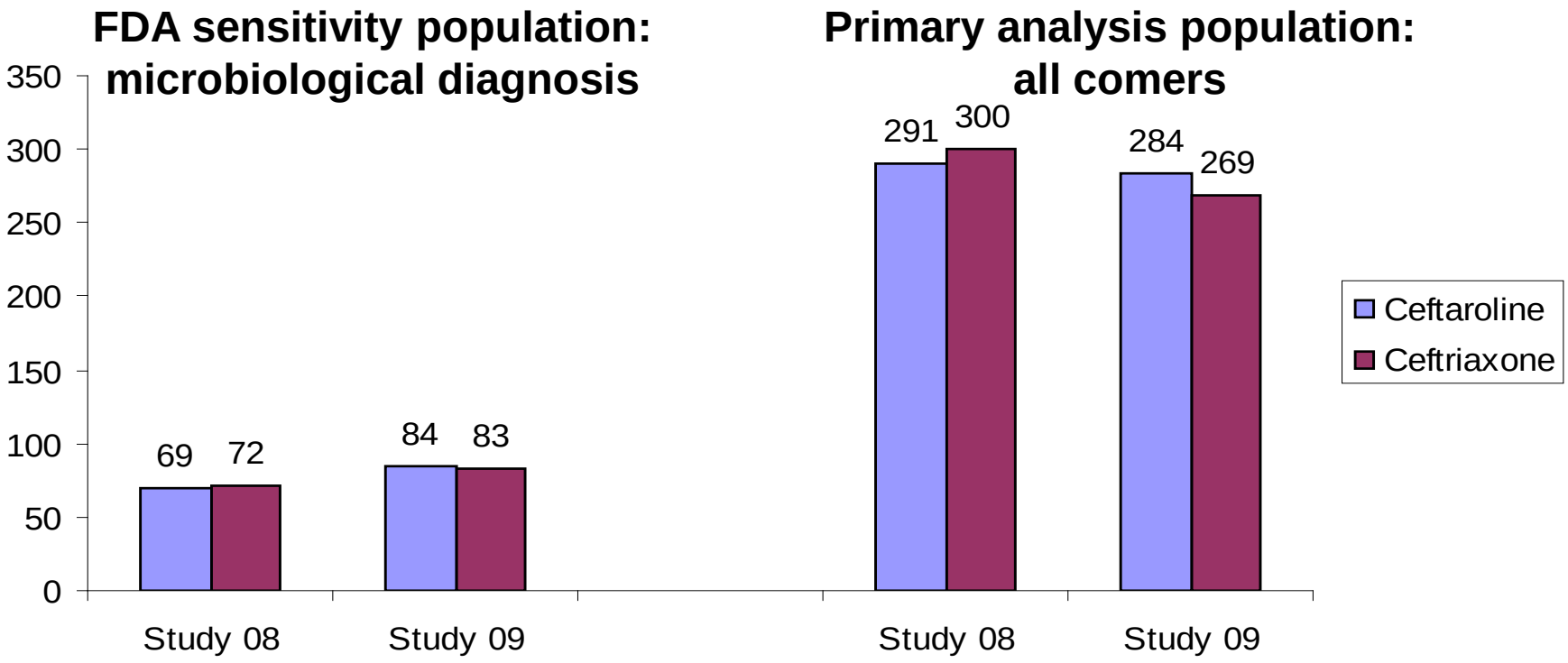
Pooled Daptomycin Trials: CE Subgroups	Clinical Response		Difference 95% CI
	Daptomycin	Ceftriaxone	
No prior effective therapy	205/272 (75.4)	245/279 (87.8)	-12.4 (-18.8, -6.0)
Prior effective therapy	88/97 (90.7)	81/92 (88.0)	2.7 (-6.1, 11.5)

Source: Pertel et al. (2008). Effects of prior effective therapy on the efficacy of daptomycin and ceftriaxone for the treatment of community-acquired pneumonia. *Clinical Infectious Diseases.*

Pooled Ceftaroline Trials: MITTE Subgroups	Clinical Response		Difference 95% CI
	Ceftaroline	Ceftriaxone	
No prior therapy	290/343 (84.5)	233/313 (74.4)	10.1 (4.0, 16.3)
Any prior therapy	185/232 (79.7)	203/256 (79.3)	0.4 (-6.8, 7.6)

Source: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2010/200327Orig1s000StatR.pdf

CABP: Microbiological Enrichment



Source: Adapted from September 7, 2010 FDA anti-infective advisory committee

CABP: Patient Severity

	Clinical Response		
	Cethromycin	Clarithromycin	Difference 95% CI
PORT Class I-II: less severe subjects	393/462 (85.1)	384/450 (85.3)	-0.3 (-4.9, 4.4)
PORT Class III-IV: more severe subjects	37/56 (66.1)	46/57 (80.7)	-14.6 (-30.5, 1.7)

Source: Adapted from June 2, 2009 FDA anti-infective advisory committee meeting

CABP: Endpoint Definition and Timing

- Traditional clinical response endpoint:
  - Resolution of signs and symptoms of the disease to the extent that additional therapy is not necessary, in the investigator's overall opinion
- Non-margin Endpoint Issues:
  - Meets regulatory standard (21CFR314.126) that adequate and well-controlled trials have a "well-defined and reliable" response assessment?
  - Composite measure of surrogates/biomarkers?
- FNIH Biomarkers Consortium endpoint:
  - Improvement with no worsening by Day 3-5 on two of the four major symptoms of cough, dyspnea, chest pain, and sputum production

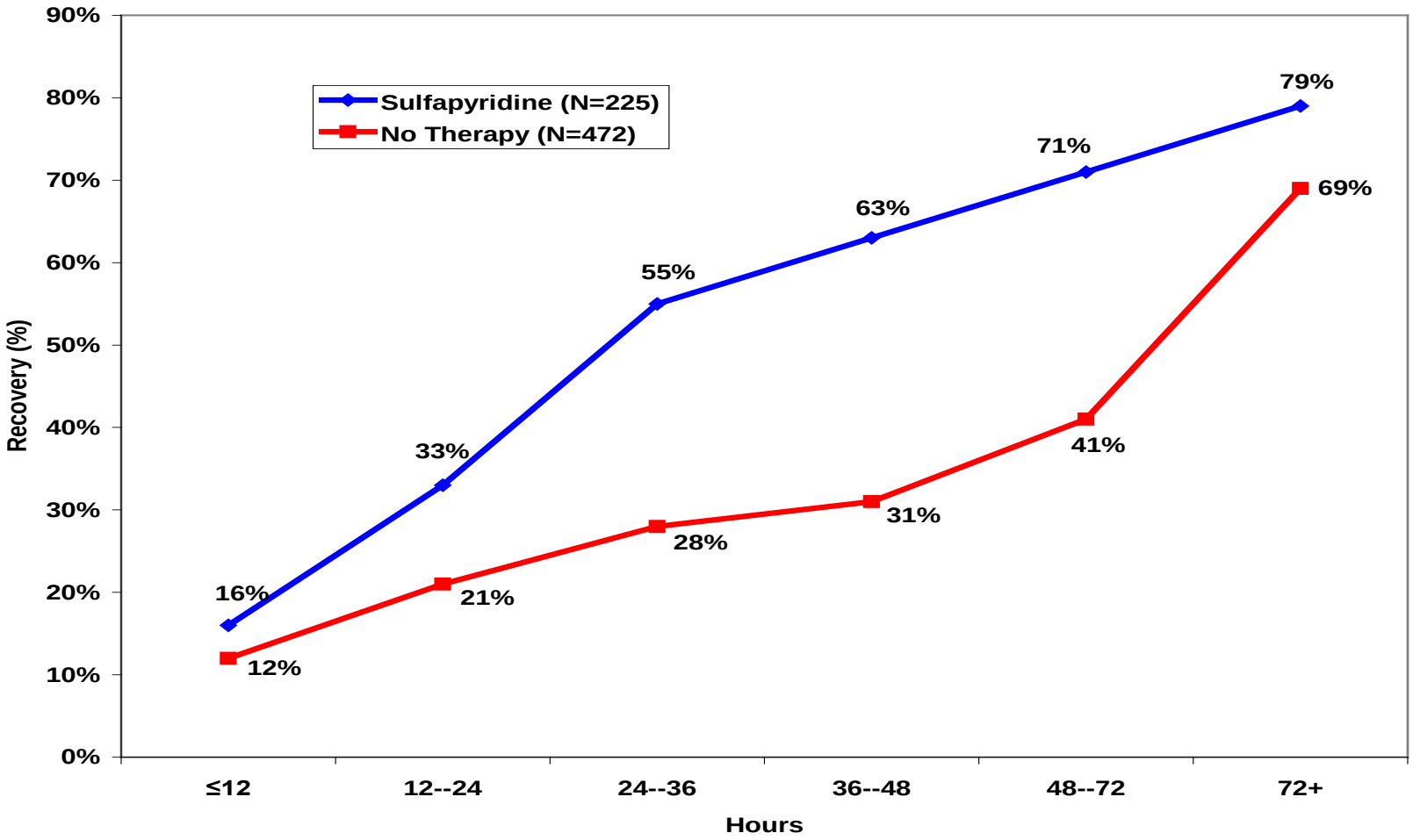
CABP: Endpoint Definition and Timing

- Natural history described in the pre-antibiotic era:
  - At first, steady deterioration and worsening respiratory symptoms
  - Recovery begins after “crisis” event (drenching sweat) around Day 8 or 9

- The drug effect seems to be for:
  - Reducing mortality
  - Preventing progression or metastatic spread of disease
  - Rapid improvement in major symptoms

- Clinical Response often defined 1-2 weeks after end of therapy

CABP: Endpoint Definition and Timing



Source: Finland et al. (1940). *Annals of Internal Medicine*.

FDA Proposals at November 2011 Anti-Infective Drugs Advisory Committee

- Prior antibacterial therapy:
  - Exclude subjects with potentially effective prior therapy
- Microbiological diagnosis:
  - Conduct two trials, require 10% NI margin in ITT of each trial, 15% NI margin in pooled subgroups with microbiologically confirmed pneumonia
- Patient severity:
  - Exclude subjects in PORT Risk Class I, allow at most 25% in Risk Class II
- Endpoint definition and timing:
  - FNIH symptom-based endpoint on Day 3-5

Questions

- **Question 1**: Please discuss design and analysis considerations that may impact the ability of a non inferiority trial to differentiate effective and ineffective therapy, such as enrollment stratification (e.g., use of prior therapy); use of a sub-ITT population (e.g., assay positive for organism) for primary efficacy analysis, handling missing data and protocol violations, etc.