Choice of outcomes in RCTs and observational studies of interventions in infectious diseases

Jesús Rodríguez Baño
UGC Intercentros E. Infecciosas, Microbiología y M. Preventiva
Hospitales Universitarios V. Macarena y V. del Rocío – IBiS - Universidad de Sevilla
Red Española de Investigación en Patología Infecciosa
• A retrospective cohort study compares the efficacy of 2 antibiotic regimens in the treatment of bloodstream infections due to carbapenemase-producer

• The main outcome is clinical cure

• Any comment?
• A randomised controlled trial compares the efficacy of an immunomodulator vs placebo in the treatment of refractory septic shock
• The main outcome is clinical length of ICU stay
• Any comment?
Why this topic?

• Studies with the inadequate outcome are frequent
• Their results are sometimes completely invalid
• Deciding the adequate outcome is tricky
Outline

• Classification or information bias
• Outcomes
  – Types
  – Selection criteria
  – Analysis
• Registration
Let’s start from the beginning...

Exposure to a variable (E)

Development of an outcome (O)
Information/classification/measurement bias

- Exposed patients are classified as unexposed or vice versa
- Patients with the outcome are classified as not having the outcome or vice versa

- Due to
  - Inadequate choice of E, O
  - Inadequate definition of E, O
  - Inadequate measuring of E, O
Some myth characters from Sevilla

- Don Juan (Don Giovanni)
- El Barbero de Sevilla
- Carmen
Some false myths about RCTs

• Observationals are biased but RCTs are not
• RCTs have a high generalizability
Bartolomé E. Murillo
Sevilla, 1617-1682

“La Inmaculada”
Museo de Bellas Artes
Terms

- Outcome, endpoint, event
- Primary, secondary
- Composite
- Proxy
- Soft, hard
Outcome definition and assessment

• Criteria: sensitive and specific enough
  – Universally accepted if possible!
• Feasible and applicable
  – Source of data (available, reliable)
• Assessment
  – Blinded vs non-blinded
  – Trained investigator
  – Monitoring
Primary outcome

• All analytical studies must have one (just one)
• Needed to calculate simple size
• Selection criteria
  – Measures the (potential) effect of the exposure
  – Is clinically relevant
  – Is frequent enough
• Would you choose as primary outcome...
  – 30-day mortality - trial on cystitis?
  – 1-year mortality - trial of therapies in naive HIV?
  – Cure rate – trial in uncomplicated SSSI infections?
  – 30-day clinical cure – trial of non-severe CAP?
  – Hospital stay for Ebola?
Dichotomous outcomes:

- Yes/no
  - Usually includes a fixed time for evaluation (or time until...)
  - One criteria or a composite
  - Proxy
- Examples
  - Infection (case-control study on risk factors)
  - Clinical cure / failure in a fixed time
  - Microbiological cure / failure in a fixed time
  - Clinical and microbiological cure / failure
  - Clinical cure + no recurrence / failure or recurrence
  - Mortality in a fixed time
  - Undetectable viral load / no
Continuous outcomes...

- Average change in a continuous parameter (SOFA...)
- Average days of... (hospital stay, antibiotics)
- Average cost
- Antibiotic consumption (DDD/1000 patient-days)
- Infection/pathogen rate
Effect of Empirical Treatment With Moxifloxacin and Meropenem vs Meropenem on Sepsis-Related Organ Dysfunction in Patients With Severe Sepsis: A Randomized Trial

**Design, Setting, and Patients** A randomized, open-label, parallel-group trial of 600 patients who fulfilled criteria for severe sepsis or septic shock (n=298 for monotherapy and n=302 for combination therapy). The trial was performed at 44 intensive care units in Germany from October 16, 2007, to March 23, 2010. The number of evaluable patients was 273 in the monotherapy group and 278 in the combination therapy group.

**Interventions** Intravenous meropenem (1 g every 8 hours) and moxifloxacin (400 mg every 24 hours) or meropenem alone. The intervention was recommended for 7 days and up to a maximum of 14 days after randomization or until discharge from the intensive care unit or death, whichever occurred first.

**Main Outcome Measure** Degree of organ failure (mean of daily total Sequential Organ Failure Assessment [SOFA] scores over 14 days; score range: 0-24 points with higher scores indicating worse organ failure); secondary outcome: 28-day and 90-day all-cause mortality. Survivors were followed up for 90 days.
Figure 2. Daily Sequential Organ Failure Assessment (SOFA) Scores

A. Daily SOFA scores in intent-to-treat analysis

B. Daily SOFA scores in per-protocol analysis

No. of evaluable patients

Meropenem alone: 249, 212, 167, 137, 124, 103, 89
Moxifloxacin and meropenem: 255, 209, 179, 153, 125, 95, 81

Meropenem alone: 181, 156, 122, 96, 88, 71, 63
Moxifloxacin and meropenem: 198, 165, 141, 119, 96, 71, 57
<table>
<thead>
<tr>
<th>Outcome</th>
<th>All Patients (N = 551)</th>
<th>Meropenem Alone (n = 273)</th>
<th>Moxifloxacin and Meropenem (n = 278)</th>
<th>P Value^a</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOFA score, mean (95% CI)</td>
<td>8.1 (7.8-8.5)</td>
<td>7.9 (7.5-8.4)</td>
<td>8.3 (7.8-8.8)</td>
<td>.36</td>
</tr>
<tr>
<td>Mortality, No. (%) [95% CI]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 28 d</td>
<td>125 (22.9) [19.5-26.7]</td>
<td>59 (21.9) [17.1-27.4]</td>
<td>66 (23.9) [19.0-29.4]</td>
<td>.58</td>
</tr>
<tr>
<td>At 90 d</td>
<td>180 (33.7) [29.7-37.9]</td>
<td>84 (32.1) [26.6-38.1]</td>
<td>96 (35.3) [29.6-41.3]</td>
<td>.43</td>
</tr>
<tr>
<td>SOFA subscores, median (IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>2.0 (1.1-3.2)</td>
<td>2.0 (1.1-3.1)</td>
<td>2.0 (1.1-3.3)</td>
<td>.54</td>
</tr>
<tr>
<td>Respiratory</td>
<td>2.4 (2.0-2.9)</td>
<td>2.4 (2.0-2.8)</td>
<td>2.5 (2.0-2.9)</td>
<td>.02</td>
</tr>
<tr>
<td>Coagulation</td>
<td>0.3 (0-1.0)</td>
<td>0.2 (0-1.0)</td>
<td>0.3 (0-1.1)</td>
<td>.48</td>
</tr>
<tr>
<td>Renal</td>
<td>0.4 (0-1.1)</td>
<td>0.3 (0-1.6)</td>
<td>0.4 (0-1.8)</td>
<td>.28</td>
</tr>
<tr>
<td>Hepatic</td>
<td>0.0 (0-0.5)</td>
<td>0 (0-0.5)</td>
<td>0 (0-0.5)</td>
<td>.86</td>
</tr>
<tr>
<td>Central nervous system</td>
<td>0.9 (0.2-2.1)</td>
<td>0.9 (0-2.3)</td>
<td>0.9 (0-2.0)</td>
<td>.65</td>
</tr>
<tr>
<td>Length of stay, median (IQR), d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In ICU</td>
<td>12 (6-22)</td>
<td>11 (5-24)</td>
<td>12 (6-21)</td>
<td>.90</td>
</tr>
<tr>
<td>In hospital</td>
<td>27 (14-44)</td>
<td>29 (14-45)</td>
<td>26 (15-42)</td>
<td>.64</td>
</tr>
<tr>
<td>Intervention-free days, median (IQR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ventilator</td>
<td>8 (0-6)</td>
<td>2 (0-6)</td>
<td>2 (0-6)</td>
<td>.59</td>
</tr>
<tr>
<td>Renal replacement therapy</td>
<td>8 (3-16)</td>
<td>8 (3-19)</td>
<td>8 (3-14)</td>
<td>.43</td>
</tr>
<tr>
<td>Antibiotic</td>
<td>0 (0-4)</td>
<td>0 (0-4)</td>
<td>0 (0-4)</td>
<td>.95</td>
</tr>
<tr>
<td>Vasopressor</td>
<td>4 (1-7)</td>
<td>3 (1-8)</td>
<td>4 (0-7)</td>
<td>.73</td>
</tr>
<tr>
<td>Secondary infection, No. (%) [95% CI]^c</td>
<td></td>
<td></td>
<td></td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>176 (32.1) [28.2-36.1]</td>
<td>89 (32.7) [27.2-38.7]</td>
<td>87 (31.4) [26.0-37.2]</td>
<td></td>
</tr>
<tr>
<td>Resistance to antibiotic, No. (%) [95% CI]^d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meropenem</td>
<td>9 (5.4) [2.5-10.0]</td>
<td>8 (9.1) [4.0-17.1]</td>
<td>1 (1.3) [0.03-6.9]</td>
<td>.04</td>
</tr>
<tr>
<td>Moxifloxacin</td>
<td>27 (26.7) [18.4-36.5]</td>
<td>16 (29.1) [17.6-42.9]</td>
<td>11 (23.9) [12.6-38.8]</td>
<td>.65</td>
</tr>
<tr>
<td>Tobramycin or gentamicin</td>
<td>16 (9.4) [5.4-14.8]</td>
<td>12 (13.3) [7.1-22.1]</td>
<td>4 (4.9) [1.4-12.2]</td>
<td>.07</td>
</tr>
</tbody>
</table>
Effect of Early versus Deferred Antiretroviral Therapy for HIV on Survival

Mari M. Kitahata, M.D., M.P.H., Stephen J. Gange, Ph.D., Alison G. Abraham, Ph.D., Barry Merriman, M.A.,
Michael S. Saag, M.D., Amy C. Justice, M.D., Ph.D., Robert S. Hogg, Ph.D., Steven G. Deeks, M.D.,
Joseph J. Eron, M.D., John T. Brooks, M.D., Sean B. Rourke, Ph.D., M. John Gill, M.B., Ch.B., Ronald J. Bosch, Ph.D.,
Jeffrey N. Martin, M.D., M.P.H., Marina B. Klein, M.D., Lisa P. Jacobson, Sc.D., Benigno Rodriguez, M.D.,
Timothy R. Sterling, M.D., Gregory D. Kirk, M.D., Ph.D., Sonia Napravnik, Ph.D., Anita P. Collier, M.D.,
Liviana M. Calzavara, Ph.D., Michael A. Horberg, M.D., Michael J. Silverberg, Ph.D., Kelly A. Gebo, M.D., M.P.H.,
James J. Goedert, M.D., Constance A. Benson, M.D., Ann C. Collier, M.D., Stephen J. Van Rompaey, Ph.D.,
Heidi M. Crane, M.D., M.P.H., Rosemary G. McKaig, Ph.D., Bryan Lau, Ph.D., Priscilla M. Freeman, M.A.,
and Richard D. Moore, M.D., for the NA-ACCORD Investigators

METHODS

We conducted two parallel analyses involving a total of 17,517 asymptomatic patients with HIV infection in the United States and Canada who received medical care during the period from 1996 through 2005. None of the patients had undergone previous antiretroviral therapy. In each group, we stratified the patients according to the CD4+ count (351 to 500 cells per cubic millimeter or >500 cells per cubic millimeter) at the initiation of antiretroviral therapy. In each group, we compared the relative risk of death for patients who initiated therapy when the CD4+ count was above each of the two thresholds of interest (early-therapy group) with that of patients who deferred therapy until the CD4+ count fell below these thresholds (deferred-therapy group).
<table>
<thead>
<tr>
<th>Variable</th>
<th>351-to-500 CD4+ Count</th>
<th></th>
<th>More-Than-500 CD4+ Count</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Relative Risk (95% CI)</td>
<td>P Value</td>
<td>Relative Risk (95% CI)</td>
<td>P Value</td>
</tr>
<tr>
<td><strong>Without inclusion of HIV RNA data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deferral of antiretroviral therapy</td>
<td>1.69 (1.26–2.26)</td>
<td>&lt;0.001</td>
<td>1.94 (1.37–2.79)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Female sex</td>
<td>1.21 (0.89–1.64)</td>
<td>0.24</td>
<td>1.85 (1.33–2.59)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Older age (per 10-yr increment)</td>
<td>1.68 (1.48–1.91)</td>
<td>&lt;0.001</td>
<td>1.83 (1.62–2.06)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Baseline CD4+ count (per 100 cells/mm³)</td>
<td>1.13 (0.72–1.78)</td>
<td>0.59</td>
<td>0.93 (0.87–0.99)</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>With inclusion of HIV RNA data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deferral of antiretroviral therapy</td>
<td>1.63 (1.21–2.19)</td>
<td>0.002</td>
<td>1.85 (1.20–2.86)</td>
<td>0.006</td>
</tr>
<tr>
<td>Female sex</td>
<td>1.47 (1.02–2.12)</td>
<td>0.04</td>
<td>1.35 (0.85–2.15)</td>
<td>0.20</td>
</tr>
<tr>
<td>Older age (per 10-year increment)</td>
<td>1.89 (1.69–2.11)</td>
<td>&lt;0.001</td>
<td>1.81 (1.58–2.07)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Baseline CD4+ count (per 100 cells/mm³)</td>
<td>0.74 (0.55–1.00)</td>
<td>0.06</td>
<td>0.97 (0.89–1.05)</td>
<td>0.45</td>
</tr>
<tr>
<td>Baseline HIV RNA level (per log₁₀ copies/ml)</td>
<td>1.11 (0.96–1.28)</td>
<td>0.15</td>
<td>1.13 (0.96–1.33)</td>
<td>0.14</td>
</tr>
</tbody>
</table>

* The CD4+ count was measured in cells per cubic millimeter. Results were calculated with the use of Cox regression analyses with inverse probability-of-censoring weights. HIV denotes human immunodeficiency virus.
Fosfomycin versus meropenem in bacteraemic urinary tract infections caused by extended-spectrum β-lactamase-producing *Escherichia coli* (FOREST): study protocol for an investigator-driven randomised controlled trial

Clara Rosso-Fernández, Jesús Sojo-Dorado, Angel Barriga, Lucía Lavin-Alconero, Zaira Palacios, Inmaculada López-Hernández, Vicente Merino, Manuel Camean, Alvaro Pascual, Jesús Rodríguez-Baño, and the FOREST Study Group

ESBL-E. coli bacteraemia
UTI*, consented patient

Randomisation

Experimental arm:
Intravenous disodium fosfomycin 4 gr/iv/6h in 60 minutes infusion
Fosfomycin trometamol orally 3 g/48 hours

Control arm:
Intravenous meropenem 1g/iv/8h in 15-30 minutes infusion
Ciprofloxacin 500 mg/12h. or Amoxicillin / clavulanate 500 mg/8h. or Trimethoprim-sulfamethoxazole or 160/800 mg/12h.

Follow-up
D3 (all +/-2 d)
D5-7: early response
D12: end of treatment
D5-7 post: test of cure
D60 (+/-10 d): end of FU

Primary outcome
Clinical/microbiological cure at D 5-7 post

Secondary outcomes
PK fosfomycin D3
Early clinical/microbiological cure at D 5-7
Length of hospital stay
Mortality up to D60
Relapse rate up to D60
Reinfection rate up to D60
Safety up to D60

Switch oral therapy after 5 days) if appropriate
- Clinical improvement
- Haemodynamic stability
- Tolerance to oral intake
- Antibiotic results
ESBL-E.coli bacteraemia UTI*, consented patient

Randomisation

Experimental arm:
Intravenous disodium fosfomycin
4 gr/iv/6h in 60 minutes infusion

Control arm:
Intravenous meropenem 1g/iv/8h
in 15-30 minutes infusion

- Clinical improvement
- Haemodynamic stability
- Tolerance to oral intake
- Antibiogram results

Fosfomycin trometamol orally
3 g /48 hours

Ciprofloxacin 500 mg/12h. or
Amoxicillin / clavulanate 500 mg/8h. or
Trimethoprim-sulfamethoxazole or
160/800 mg/12h.

Follow-up
D3 (all +/- 2 d)
D5-7: early response
D12: end of treatment
D5-7 post: test of cure
D60 (+/- 10 d): end of FU

Primary outcome
Clinical/microbiological cure at D 5-7 post

Blinded external assessment

Secondary outcomes
PK fosfomycin D3
Early clinical/microbiological cure at D 5-7
Length of hospital stay
Mortality up to D60
Relapse rate up to D60
Reinfection rate up to D60
Safety up to D60
• For comparison of efficacy of antiretroviral drugs in naive patients, the proportion of patients with undetectable viral load at fixed times is used as outcome.

• Any comment?
Diego Velázquez
Sevilla, 1599-1660
“Santa Rufina”
Fundación Focus Abengoa
Hard vs soft

• Hard = “objective”
• Soft = “subjective”

• Examples
  – Hard: all-cause mortality, hospital stay, viral load
  – Soft: clinical cure, infection-related mortality
Hard vs soft

• More important if assessment is not blinded
  – Biased interpretation of definition/criteria for soft outcomes

• Why soft outcomes are sometimes needed?
  – Hard are not relevant or frequent enough
Soft outcomes: bias control

• Blind assessment
• If not, use external blind assessment + monitoring
  – Collection of objective data supporting the decisión
  – Chart review
Analysis of dichotomous outcome

• Comparison of proportions
  – Chi squared/Fisher, logistic regression
• Survival analysis (time until the outcome)
  – Kaplan-Meier curves, log rank test
  – Cox regression
• Event (yes/no) does not consider if events occur early or late
• Solution?
  – Time to event
  – Each patient sum up the number of days from exposure to event (o time of censorship)
Gentamicin therapy for sepsis due to carbapenem-resistant and colistin-resistant *Klebsiella pneumoniae*

Marcelino Gonzalez-Padilla, Julián Torre-Cisneros, Francisco Rivera-Espin, Antonio Pontes-Moreno, Lorena López-Cerero, Alvaro Pascual, Clara Natera, Marina Rodríguez, Inmaculada Salcedo, Fernando Rodriguez-López, Antonio Rivera and Jesús Rodríguez-Baño

**Figure 2.** Kaplan–Meier curves showing the impact of treatment with suboptimal targeted treatment, optimal targeted treatment without gentamicin and optimal targeted treatment with gentamicin on survival at 30 days in patients with severe infection caused by carbapenem-resistant and colistin-resistant *K. pneumoniae* (log-rank test 17.3, *P* < 0.001).
Staphylococcus aureus bloodstream infection: A pooled analysis of five prospective, observational studies

Outcome variable: time until death or censorship
Analysis of continuous outcomes...

- Student T test, Mann-Whitney U test, ANOVA
- Linear regression
Analysis criteria

• Intention-to-treat
  – All randomised patients are evaluated for the outcome
  – If changed, consider the reasons
    • Failure? All reasons?

• Per protocol
  – Only patients receiving a minimum amount of exposure are evaluated for the outcome
Fixed time assessment: mortality in bacteraemia as an example

• Dead may be related to
  – Underlying risk (chronic and acute conditions, age)
  – Direct effect of infection and related complications
  – Indirect effect of infection
What is the best time frame?

- Directly related to infection
- Related to age, chronic conditions
• What is your hypothesis/question
  – ... very early mortality? (day 7-14)
  – ... early mortality? (day 30)
  – ... late mortality? (day 90)
  – ... very late mortality? (1 year)
A Systematic Review of the Methods Used to Assess the Association between Appropriate Antibiotic Therapy and Mortality in Bacteremic Patients

Jessina C. McGregor,1 Shayna E. Rich,2 Anthony D. Harris,2,4 Eli N. Perencevich,2,4 Regina Obar,5 Thomas P. Lodise, Jr.,5 Ram R. Miller,2 and Jon P. Furuno2

Clinical Infectious Diseases 2007;45:329–37

Mortality should be measured in a manner that best represents the underlying construct within the biologically plausible window of effect. Statistical analyses should be used to account for loss to follow-up (e.g., because of hospital discharge).
Gonzalo Bilbao
Sevilla, 1860-1038

“Las cigarreras”
Museo de Bellas Artes
Desirability of Outcome Ranking (DOQR) and Response Adjusted for Duration of Antibiotic Risk (RADAR)

Scott R. Evans,1 Daniel Rubin,2 Dean Follmann,3 Gene Pennello,4 W. Charles Huskins,5 John H. Powers,6,7 David Schoenfeld,8 Christy Chuang-Stein,9 Sara E. Cosgrove,10 Vance G. Fowler Jr,11 Ebbing Lautenbach,12 and Henry F. Chambers13

Clinical Infectious Diseases® 2015;61(5):800–6

IBiS INSTITUTO BIOMEDICINA DE SEVILLA
reipi-isciii

Hospital Universitario Virgen Macarena, Sevilla
ESCMID COLLABORATIVE CENTRE
Challenges in design of studies to evaluate antimicrobial use strategies

• Misleading outcomes
  – Hospital stay (lower if patient dies)
  – Antibiotic use per days (lower if higher hospital stay)

• Non-inferiority issues
  – Not attractive to patients (unethical?)
  – Large simple size

• Lack of integration of benefits and harms
  – Usually analysed separately...
  – ...but what is the net effect?
Proposal

• Desirability of outcome ranking (DOOR)
• Response adjusted for duration of antibiotic use risk (RADAR)
• Phylosophy: use each patient net experience
• Process
  – Define the clinical important outcomes
  – Define an ordinal composite
• Based on clinical relevance and patients preferences
Table 1. Suggestions for Overall Clinical Outcome Construction

- Consider:
  (a) The general overall (benefit and risk) patient-level clinical outcomes that have differing levels of importance, and
  (b) How patients tend to cluster themselves in terms of overall clinical outcomes (i.e., categories may be naturally apparent).
- Identify and prioritize important clinical outcomes including efficacy, safety, and quality of life. Some factors can be viewed as equally important.
- Use “all-cause” outcomes when formulating the response in randomized studies. Patients, not specific outcomes, are being evaluated. Causality is evaluated by a contrast of the randomized strategies rather than judged “relatedness” to the intervention or disease. If an outcome is unrelated to treatment, then it will occur with similar frequency between randomized arms. If an outcome occurs differentially between arms, then it is related to treatment.
- Consider using outcomes that are a function of the patient, or standardize the criteria for clinical decisions (e.g., duration of hospital stay), to eliminate or reduce variation induced by clinician decision. A fundamental tenet in clinical trials is to minimize variation as this provides the best opportunity to identify intervention effects if they exist. Although some measures (e.g., change of therapy) can be objectively measured, they are partly a function of clinician decision in addition to patient outcome. This adds another source of variation (e.g., due to the clinician). Patients may switch therapy because of clinical failure but should not necessarily be considered a clinical failure because they switch therapy. Clinician decisions can be used as a surrogate for patient responses when detailed patient responses are unavailable.
- Use endpoints that are clinically meaningful and that are measures of how patients feel, function, or survive. Limit the use of biomarkers unless they are thoroughly validated as surrogates for clinical patient response. Many currently utilized biomarkers have not been validated as adequate surrogates. If patient responses are observable within reasonable time frames, then surrogates may be unnecessary.
• An example
• Outcomes: cure, survival, adverse events
• Ordinal composite
  1. Cure, survival, no AE
  2. Cure, survival, some mild AE
  3. No cure, survival, no AE
  4. No cure, survival, AE
  5. Death
CAP in children: short (5 d) vs standard (10 d) beta-lactam therapy

<table>
<thead>
<tr>
<th>Rank</th>
<th>Outcome Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Survival; adequate clinical response; no adverse events</td>
</tr>
<tr>
<td>2.</td>
<td>Survival; adequate clinical response; mild adverse event(s)</td>
</tr>
<tr>
<td>3.</td>
<td>Survival; adequate clinical response; moderate adverse event(s)</td>
</tr>
<tr>
<td>4.</td>
<td>Survival; adequate clinical response; severe adverse event(s)</td>
</tr>
<tr>
<td>5.</td>
<td>Survival; inadequate clinical response without additional emergency department or clinic visit or hospitalization</td>
</tr>
<tr>
<td>6.</td>
<td>Survival; inadequate clinical response with additional emergency department or clinic visit but without hospitalization; any grade of adverse event</td>
</tr>
<tr>
<td>7.</td>
<td>Survival; inadequate clinical response with hospitalization; any grade of adverse event</td>
</tr>
<tr>
<td>8.</td>
<td>Death</td>
</tr>
</tbody>
</table>
RADAR

• A version of DOOR for superiority trials
• Outcomes (as in DOOR) + duration of therapy (less is better, but not at the expense of clinical outcomes)
• Process
  – First, clinical outcomes are evaluated and ranked
  – Then, duration of therapy; for the same clinical outcomes, shorter duration of therapy ranks higher
Table 2. Response Adjusted for Duration of Antibiotic Risk Illustration: Participant Data and Data Summaries

<table>
<thead>
<tr>
<th>Participant</th>
<th>Treatment Arm</th>
<th>Overall Clinical Outcome(^a)</th>
<th>Days of Antibiotic Use</th>
<th>DOOR</th>
<th>No. of Control Participants (n = 13) With a Lower DOOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>New</td>
<td>2</td>
<td>5</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>New</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>C</td>
<td>New</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>D</td>
<td>New</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>E</td>
<td>New</td>
<td>3</td>
<td>3</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>F</td>
<td>New</td>
<td>2</td>
<td>3</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>G</td>
<td>New</td>
<td>3</td>
<td>5</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>H</td>
<td>New</td>
<td>3</td>
<td>2</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>I</td>
<td>New</td>
<td>1</td>
<td>5</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>New</td>
<td>3</td>
<td>8</td>
<td>23</td>
<td>3</td>
</tr>
<tr>
<td>K</td>
<td>New</td>
<td>2</td>
<td>6</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>L</td>
<td>New</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>M</td>
<td>New</td>
<td>1</td>
<td>8</td>
<td>14.5</td>
<td>7.5</td>
</tr>
<tr>
<td>N</td>
<td>Control</td>
<td>2</td>
<td>12</td>
<td>26</td>
<td>Sum = 109.5</td>
</tr>
<tr>
<td>O</td>
<td>Control</td>
<td>2</td>
<td>7</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>Control</td>
<td>1</td>
<td>9</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>Control</td>
<td>2</td>
<td>8</td>
<td>14.5</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>Control</td>
<td>3</td>
<td>6</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Control</td>
<td>2</td>
<td>11</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>Control</td>
<td>1</td>
<td>10</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>Control</td>
<td>2</td>
<td>9</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>Control</td>
<td>3</td>
<td>9</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>Control</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>Control</td>
<td>1</td>
<td>8</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>Control</td>
<td>2</td>
<td>10</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>Control</td>
<td>3</td>
<td>10</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

The probability of a better DOOR for a randomly selected participant from the new strategy compared with the old strategy is the number of between-treatment pairwise comparisons in which the new treatment has a higher DOOR than the control (109.5), divided by the total number of possible pairwise comparisons (169), resulting in 64.8% (95% confidence interval, 57%–71%).

Abbreviations: AE, adverse effects; DOOR, desirability of outcome ranking.

\(^a\) Overall clinical outcome coding: 1, success without AE; 2, success with AE; 3, failure.
Statistics

• Probability of a better DOOR
  – Sum of the number of between-treatment pairwise comparison in which the experimental arm has a higher DOOR than the control divided by the total number of possible pairwise comparisons
  – Calculate 95% CI
  – 50%: no difference

• Sample size calculation
  – Null: no difference in DOOR. Alternative: >50%
  – Magnitude of superiority to detect based on minimum clinically important difference
CAP in children: short (5 d) vs standard (10 d) beta-lactam therapy

Initial sample size for non-inferiority: 800 (400 per arm)
After RADAR: superiority >60%, 360 (180 per arm)
My concerns

• Ordinal variable used as continuous for stratification
• What if one part of the composite is more important than other
  – Sensitivity analysis is proposed
• Subjectivity of ranking definition
Registration of observational studies
The next step towards research transparency

Rationale for registration of clinical trials

Ethical
Respect the investigator-participant covenant to contribute to biomedical knowledge by making trial methods and results public.
Provide global open access to information.
Reduce unnecessary duplication of invested research resources through awareness of existing trials.
Assure accountability with regard to global standards for ethical research.
Enable monitoring of adherence to ethical principles and processes.

Scientific
Increase the reliability and availability of evidence on which healthcare decisions are based.
Improve trial participation.
Increase opportunities for collaboration.
Ensure transparency of trial design and methods.
Provide open review of protocols to improve trial quality and refine methods.
Provide means for identification and prevention of biased under-reporting or over-reporting of research.
Accelerate knowledge creation.

BMJ 2010;340:c950

Most of these apply also for observational studies!!
Today's Random Medical News

From the New England Journal of Panic-Inducing Gobbledygook

According to a report released today...

Vandenbroucke, Lancet 2004
The Value of Statistical Analysis Plans in Observational Research
Defining High-Quality Research From the Start

Laine Thomas, PhD
Eric D. Peterson, MD, MPH

JAMA, August 22/29, 2012—Vol 308, No. 8 773
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration

Jan P. Vandenbroucke¹, Erik von Elm²,³, Douglas G. Altman⁴, Peter G. Gotzsche⁵, Cynthia D. Mulrow⁶, Stuart J. Pocock⁷, Charles Poole⁸, James J. Schlesselman⁹, Matthias Egger²,¹⁰* for the STROBE Initiative

PLoS Medicine 2007
Box 1. Summary of recommendations to improve the quality of observational studies for assessing the efficacy and safety of different treatment options for multidrug-resistant organism infections.

- Clearly define the population under study and use strict, clinically sound inclusion and exclusion criteria
- Clearly define criteria for assignment of patients to each treatment group. Use clinically sound criteria considering an adequate starting period, a minimum duration, and avoiding the noise caused by combined or sequential therapies
- Define outcome variables so that the effect of therapy can be adequately assessed according to the type of infection and accepted definitions
- Include all potential confounders as covariates. Provide adequate definitions for them
- Whenever possible, use a prospective and multicenter design
- Assure an adequate quality in the collection of data and, if possible, use quality monitoring
- If multiple centers are participating, be sure that the definition of variables are understood by all participants and interpreted in the same way
- Register the project in advance on a public website, including predefined hypothesis, objectives and methods, and a detailed predefined statistical analysis plan
- Use adequate statistical methods to control for confounding
Carmen Laffón
Sevilla, 1934
Conclusions

• Outcomes:
  – Appropriate choice (relevant, measures the effect, feasible)
  – Appropriate definition and criteria
  – Appropriate, unbiased assessment
  – Appropriate analysis

• Pre-register!
REVIEW

For reprint orders, please contact: reprints@futuremedicine.com

Improved treatment of multidrug-resistant bacterial infections: utility of clinical studies

Esther Bettiol1, Wouter C Rottier2, Maria Dolores del Toro3,4, Stephan Harbarth1, Marc J Bonten1,5,6 & Jesús Rodriguez-Bañó3,4; on behalf of the COMBACTE consortium
Addendum to the guideline on the evaluation of medicinal products indicated for treatment of bacterial infections.

Complicated Urinary Tract Infections: Developing Drugs for Treatment Guidance for Industry

Guidance for Industry
Evaluating the Safety of Antimicrobial New Animal Drugs with Regard to Their Microbiological Effects on Bacteria of Human Health Concern

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

February 2015
Clinical/Antimicrobial
Progress on Developing Endpoints for Registrational Clinical Trials of Community-Acquired Bacterial Pneumonia and Acute Bacterial Skin and Skin Structure Infections: Update From the Biomarkers Consortium of the Foundation for the National Institutes of Health

George H. Talbot,1 John H. Powers,2 Thomas R. Fleming,4 Judith A. Siuciak,3 John Bradley,5,6 Helen Boucher,7,8 on behalf of the CABP-ABSSSI Project Team

Clinical Infectious Diseases 2012;55(8):1114–21
Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR)

Scott R. Evans,1 Daniel Rubin,2 Dean Follmann,3 Gene Pennello,4 W. Charles Huskins,5 John H. Powers,6,7
David Schoenfeld,8 Christy Chuang-Stein,9 Sara E. Cosgrove,10 Vance G. Fowler Jr.,11 Ebbing Lautenbach,12 and
Henry F. Chambers13

Clinical Infectious Diseases® 2015;61(5):800–6