Essentials of Clinical Research:

Choosing a Study Design

David J. Cohen, M.D., M.Sc.

Director of Cardiovascular Reseach
Saint Luke's Mid America Heart Institute

University of Missouri- Kansas City
Kansas City, MO USA

Designing Clinical Research

- Randomized trials vs. observational studies
- Superiority vs. Non-inferiority trials
- Blinding and placebos

Designing Clinical Research

- Randomized trials vs. observational studies
- Superiority vs. Non-inferiority trials
- Blinding and placebos

Study Designs for Clinical Research

Weakest evid

- Single case report (anecdote)

<u>Challenge of Clinical Research</u>:

To match each clinical question to the study design that will allow it to be answered in a practical, timely, and efficient manner

study

trols

- Single randomized clinical trial
- Multiple large, randomized clinical trials

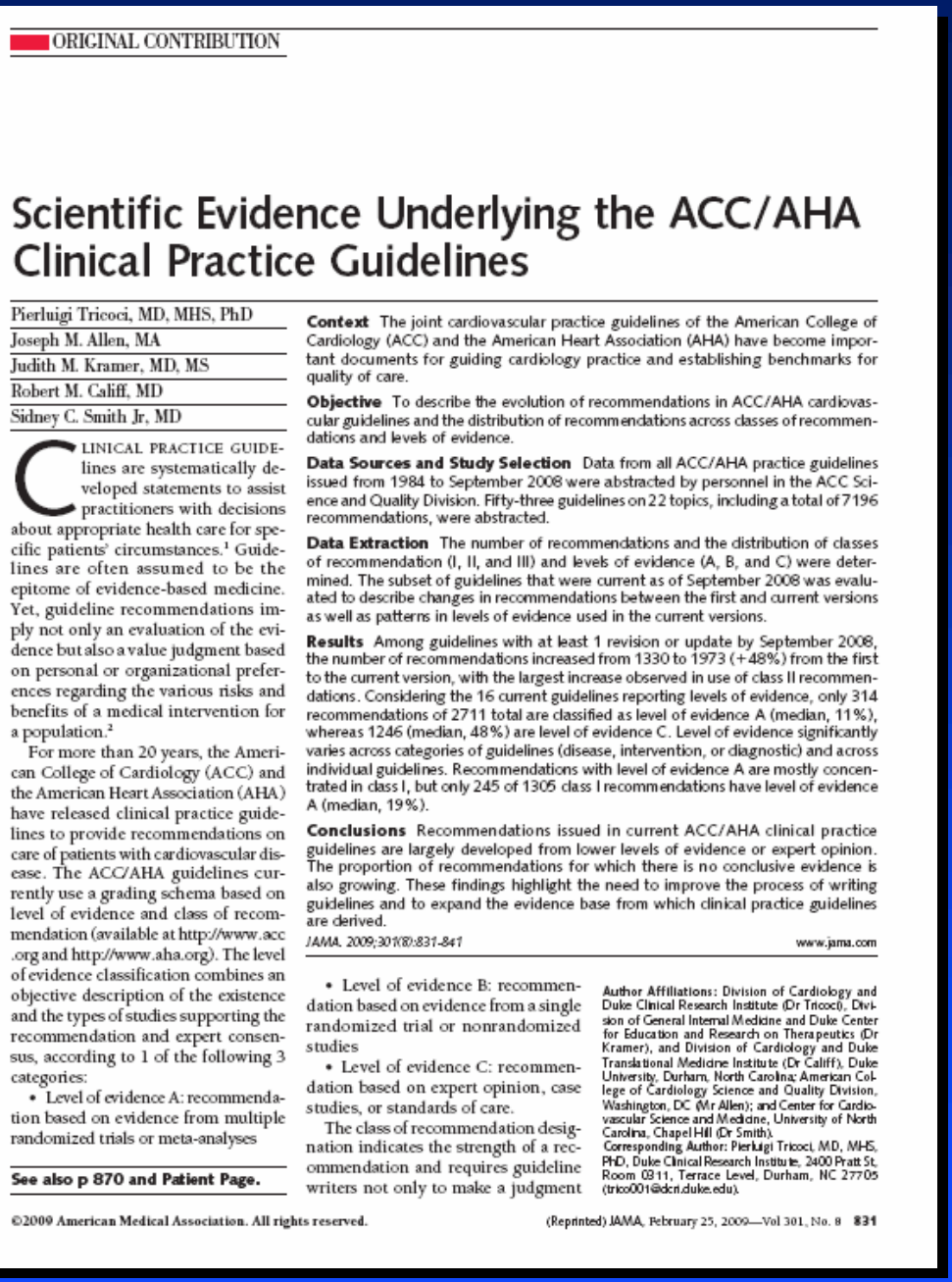
Strongest evidence

Why do we need RCTs?

- RCTs are the best available technique for eliminating bias in the assessment of a treatment effect
  - *Eliminates both measured and unmeasured confounding*
- With continued improvement in medical care, most treatment effects of interest in cardiovascular dz have only modest effects (RR reductions ~15-20%)
  - *Only RCTs can provide sufficient precision and confidence to reliably detect small benefits*
  - *Increasing emphasis on "large, simple trials" (>20K pts)*

Limitations of Clinical Trials

Only a finite # of clinical trials can be performed. Frequently, trial results may not apply to the particular patient or clinical situation in question



ORIGINAL CONTRIBUTION

Scientific Evidence Underlying the ACC/AHA Clinical Practice Guidelines

Pierluigi Tricoci, MD, MHS, PhD
Joseph M. Allen, MA
Judith M. Kramer, MD, MS
Robert M. Califf, MD
Sidney C. Smith Jr, MD

Context The joint cardiovascular practice guidelines of the American College of Cardiology (ACC) and the American Heart Association (AHA) have become important documents for guiding cardiology practice and establishing benchmarks for quality of care.

Objective To describe the evolution of recommendations in ACC/AHA cardiovascular guidelines and the distribution of recommendations across classes of recommendations and levels of evidence.

Data Sources and Study Selection Data from all ACC/AHA practice guidelines issued from 1984 to September 2008 were abstracted by personnel in the ACC Science and Quality Division. Fifty-three guidelines on 22 topics, including a total of 7196 recommendations, were abstracted.

Data Extraction The number of recommendations and the distribution of classes of recommendation (I, II, and III) and levels of evidence (A, B, and C) were determined. The subset of guidelines that were current as of September 2008 was evaluated to describe changes in recommendations between the first and current versions as well as patterns in levels of evidence used in the current versions.

Results Among guidelines with at least 1 revision or update by September 2008, the number of recommendations increased from 1330 to 1973 (+48%) from the first to the current version, with the largest increase observed in use of class II recommendations. Considering the 16 current guidelines reporting levels of evidence, only 314 recommendations of 2711 total are classified as level of evidence A (median, 11%), whereas 1246 (median, 48%) are level of evidence C. Level of evidence significantly varies across categories of guidelines (disease, intervention, or diagnostic) and across individual guidelines. Recommendations with level of evidence A are mostly concentrated in class I, but only 245 of 1305 class I recommendations have level of evidence A (median, 19%).

Conclusions Recommendations issued in current ACC/AHA clinical practice guidelines are largely developed from lower levels of evidence or expert opinion. The proportion of recommendations for which there is no conclusive evidence is also growing. These findings highlight the need to improve the process of writing guidelines and to expand the evidence base from which clinical practice guidelines are derived.

JAMA. 2009;301(8):831-841 www.jama.com

CLINICAL PRACTICE GUIDELINES are systematically developed statements to assist practitioners with decisions about appropriate health care for specific patients' circumstances.[1] Guidelines are often assumed to be the epitome of evidence-based medicine. Yet, guideline recommendations imply not only an evaluation of the evidence but also a value judgment based on personal or organizational preferences regarding the various risks and benefits of a medical intervention for a population.[2]

For more than 20 years, the American College of Cardiology (ACC) and the American Heart Association (AHA) have released clinical practice guidelines to provide recommendations on care of patients with cardiovascular disease. The ACC/AHA guidelines currently use a grading schema based on level of evidence and class of recommendation (available at http://www.acc.org and http://www.aha.org). The level of evidence classification combines an objective description of the existence and the types of studies supporting the recommendation and expert consensus, according to 1 of the following 3 categories:

- Level of evidence A: recommendation based on evidence from multiple randomized trials or meta-analyses
- Level of evidence B: recommendation based on evidence from a single randomized trial or nonrandomized studies
- Level of evidence C: recommendation based on expert opinion, case studies, or standards of care.

The class of recommendation designation indicates the strength of a recommendation and requires guideline writers not only to make a judgment

See also p 870 and Patient Page.

Author Affiliations: Division of Cardiology and Duke Clinical Research Institute (Dr Tricoci), Division of General Internal Medicine and Duke Center for Education and Research on Therapeutics (Dr Kramer), and Division of Cardiology and Duke Translational Medicine Institute (Dr Califf), Duke University, Durham, North Carolina; American College of Cardiology Science and Quality Division, Washington, DC (Mr Allen); and Center for Cardiovascular Science and Medicine, University of North Carolina, Chapel Hill (Dr Smith).
Corresponding Author: Pierluigi Tricoci, MD, MHS, PhD, Duke Clinical Research Institute, 2400 Pratt St, Room 0311, Terrace Level, Durham, NC 27705 (trico001@dcri.duke.edu).

- Reviewed all ACC/AHA practice guidelines from 1984-2008 (n=53 guidelines, 7196 recommendations)
- Levels of evidence in current guidelines
  - A (multiple RCTs)– 11%
  - B (single RCT or non- randomized studies only)– 41%
  - C (expert opinion or std of care)– 48%

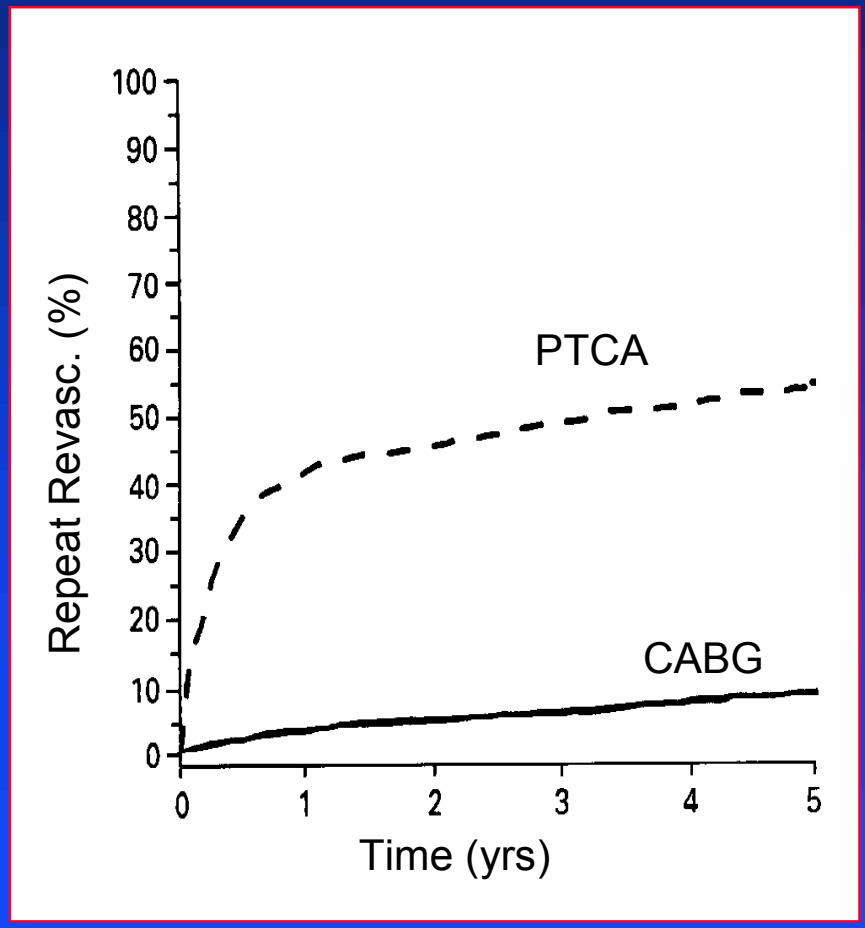
Trioci P, et al. JAMA 2009;301:831-41

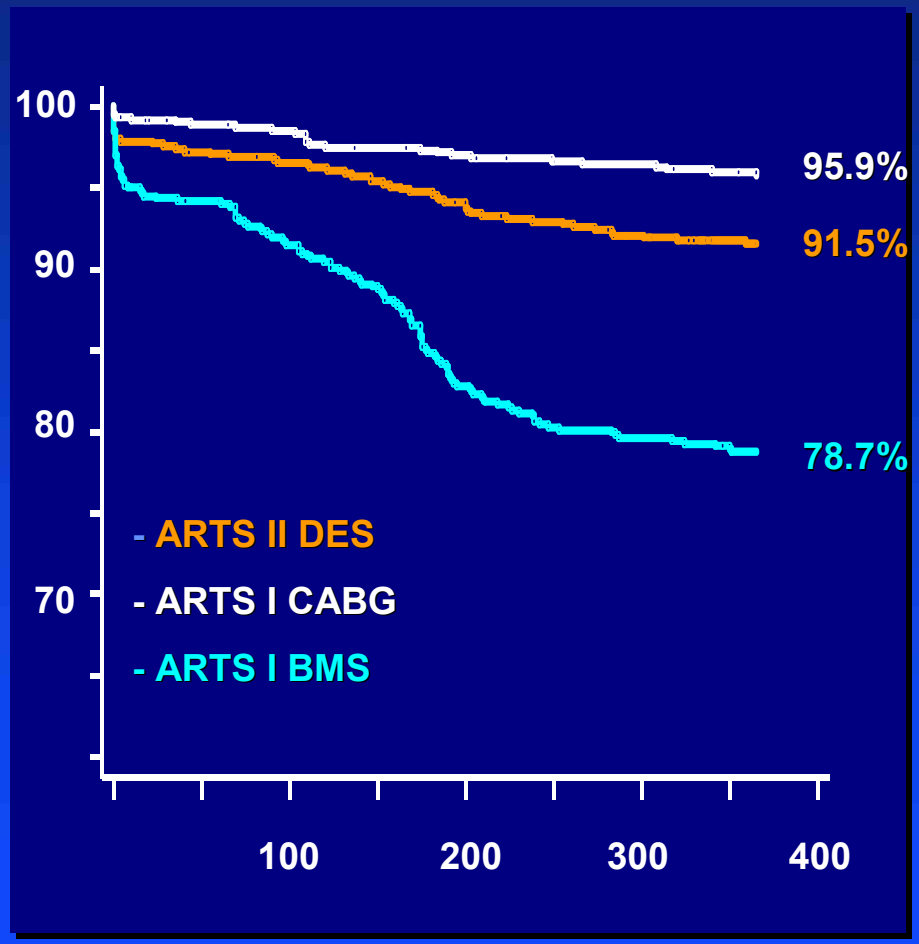
Limitations of Clinical Trials

Obsolescence

- RCT's are best suited to evaluation of "mature" therapies
- Clinical trials are a poor way to evaluate rapidly changing technologies and standards of care→ particularly problematic for medical devices
- Trials are particularly vulnerable when enrollment is slow or the follow-up duration is long

BARI: Repeat Revascularization



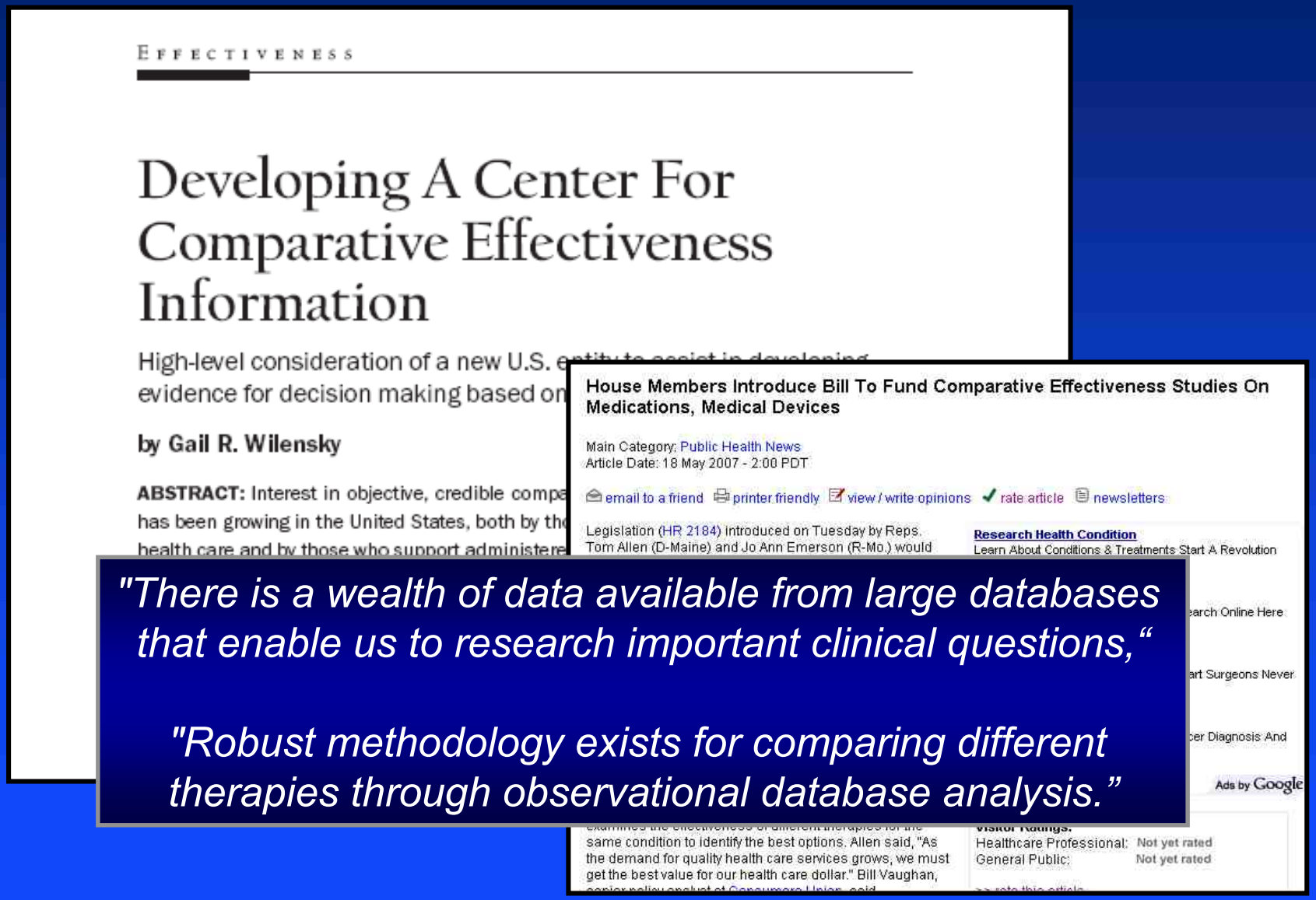


Additional Limitations of RCTs

- Often underpowered for modest treatment effects
  - *Still relevant from public health standpoint if affected population is large*
- Surrogate endpoints→ ? Clinical relevance
- Generalizability?
  - *Tend to study generally healthy patients*
  - *Treated with standardized protocols*
  - *By experienced providers*
- Certain questions not easily subject to RCT
  - *Unethical, impractical, no business case, or*
  - *Studies of harmful effects*

Can we use observational studies (registries) for clinical evidence development?

Comparative Effectiveness



EFFECTIVENESS

Developing A Center For Comparative Effectiveness Information

High-level consideration of a new U.S. entity to assist in developing evidence for decision making based on

by Gail R. Wilensky

ABSTRACT: Interest in objective, credible compa has been growing in the United States, both by th health care and by those who support administere

House Members Introduce Bill To Fund Comparative Effectiveness Studies On Medications, Medical Devices

Main Category: Public Health News
Article Date: 18 May 2007 - 2:00 PDT

email to a friend | printer friendly | view / write opinions | rate article | newsletters

Legislation (HR 2184) introduced on Tuesday by Reps. Tom Allen (D-Maine) and Jo Ann Emerson (R-Mo.) would

Research Health Condition
Learn About Conditions & Treatments Start A Revolution

earch Online Here

art Surgeons Never

cer Diagnosis And

Ads by Google

same condition to identify the best options. Allen said, "As the demand for quality health care services grows, we must get the best value for our health care dollar." Bill Vaughan,

Visitor Ratings:
Healthcare Professional: Not yet rated
General Public: Not yet rated

"There is a wealth of data available from large databases that enable us to research important clinical questions,"

"Robust methodology exists for comparing different therapies through observational database analysis."

Wilensky G Health Affairs Nov 2006:w572-w588

Registry Studies: Key Advantages

- Allows for *rapid enrollment* of large numbers of patients → accomodates changes in practice over time
- Broad inclusion criteria ensure that study's findings may be *applicable to most patients*
- Ideal for determining optimal procedural technique as well as for identifying appropriate patient subsets for treatment

Registry Studies: Key Disadvantages

Data quality and completeness

- Analysis results only as solid as the data ("Bad data in…")
- Particularly challenging with administrative datasets
- Incomplete data → rarely missing at random
- Not necessarily related to registry design, but more related to degree of rigor employed in data collection

Treatment selection bias

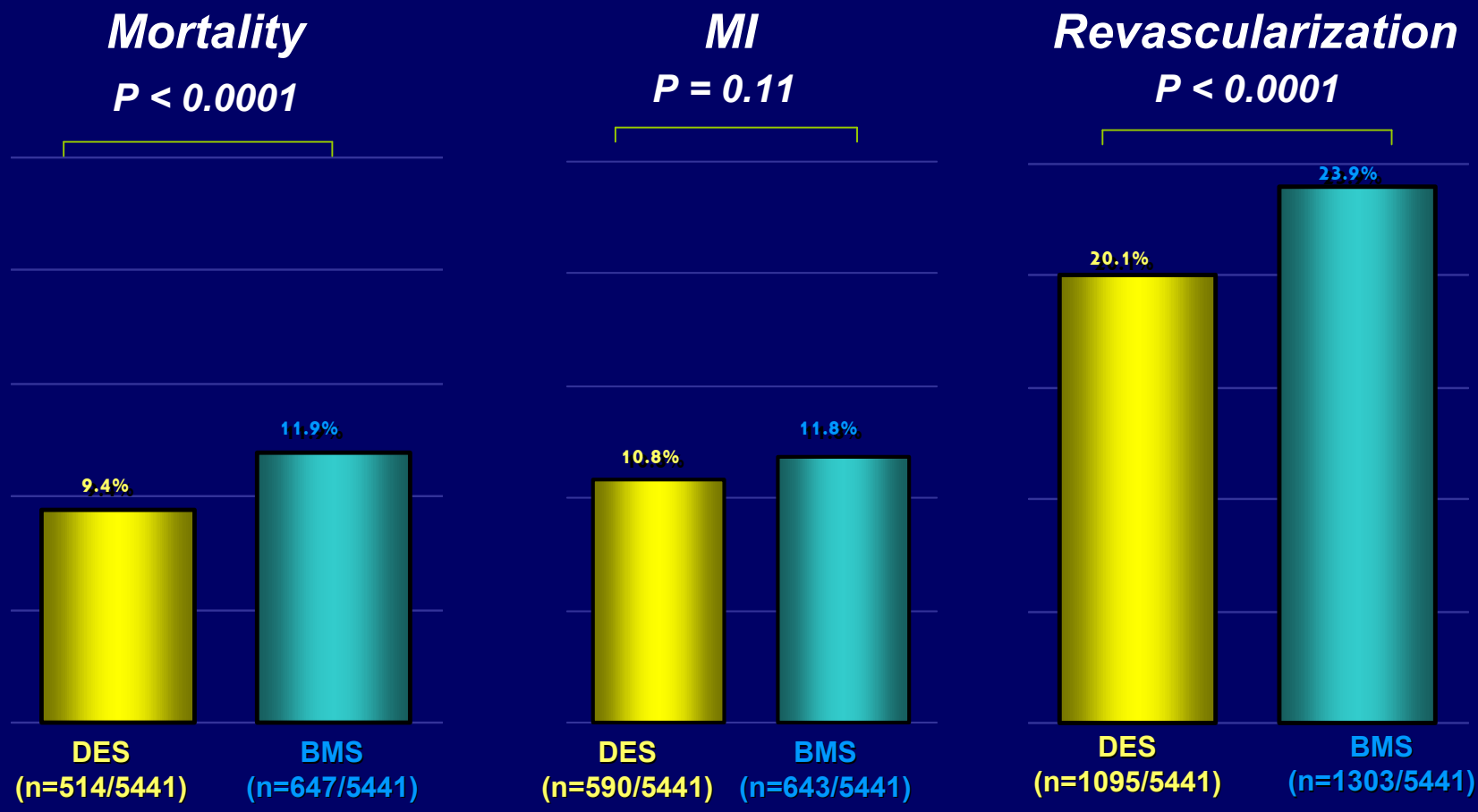
- Pt Level: risk factors, disease severity, comorbidity
- MD level: those selecting a specific treatment may differ in care process and quality
- Site-level: structural and quality of care differences

Techniques for Overcoming Selection Bias

- Regression modeling
  - *Adjust results directly for 'confounding factors' associated with treatment and outcome*
- Propensity adjustment
  - *Identify factors associated with treatment selection*
  - *Then adjust for the probability of treatment (propensity score) or match patients for this factor*
- Newer approaches
  - *Instrumental variables analysis*

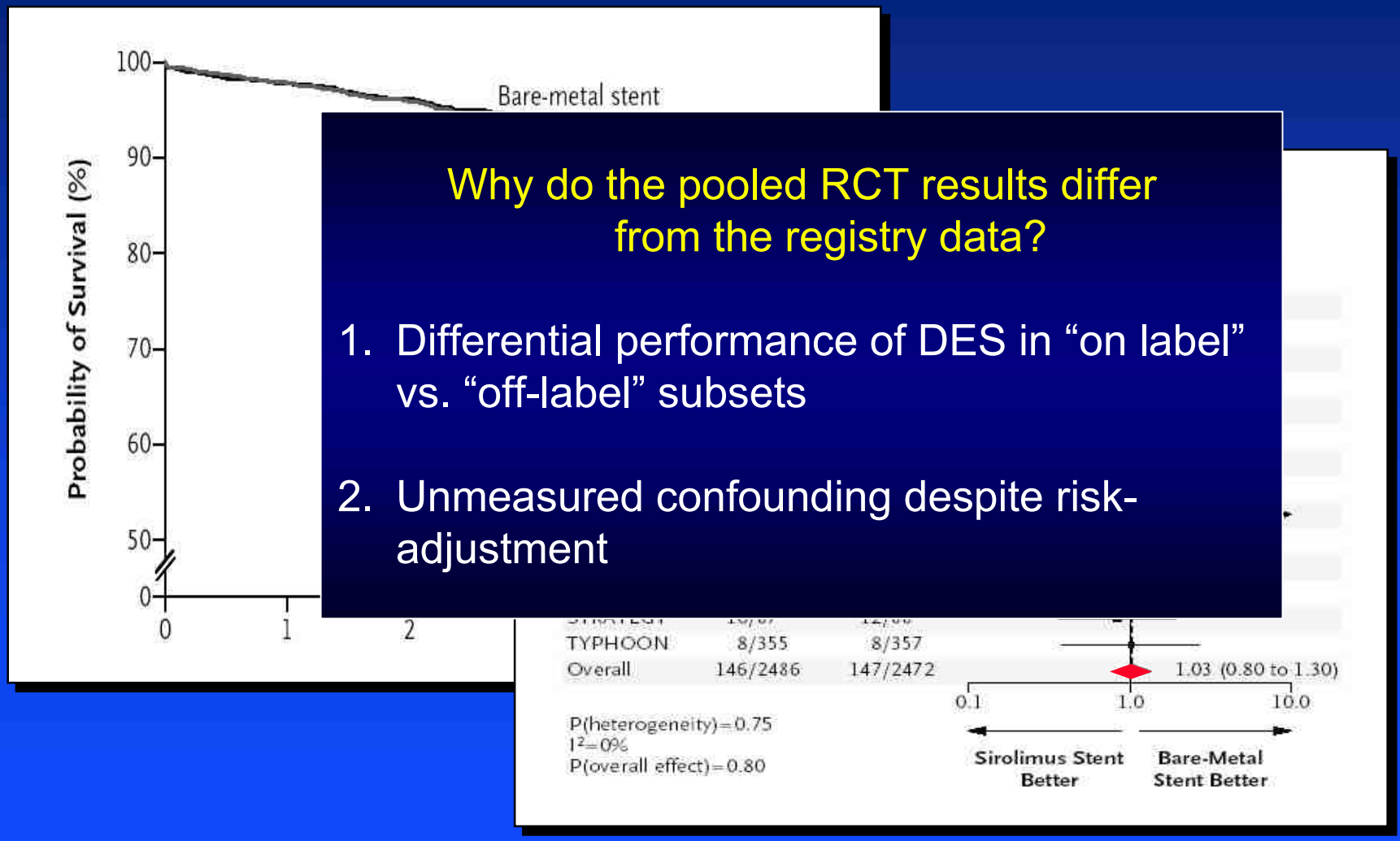
Drug-Eluting and Bare Metal Stenting in Massachusetts, Primary Results

Propensity Matched 2-Year Outcomes



Mauri L, et al. Circulation 2008;118:1817-27

Do Drug-Eluting Stents Save Lives?
Pooled RCT Results



Why do the pooled RCT results differ from the registry data?

1. Differential performance of DES in “on label” vs. “off-label” subsets
2. Unmeasured confounding despite risk-adjustment

Kastrati et al. NEJM 2007; 356:1020-9

Summary: RCTs vs. Registries

- If randomization an option, it is still by far the best and most definitive approach to developing unbiased, reliable evidence
- Nonetheless, gaps will continue to exist in our evidence base
  - *No trials*
  - *Non-representativeness (lack of generalizability)*
  - *Artificial nature of trial protocol (e.g., angiographic f/u)*
- With careful planning and analysis, observational treatment comparisons can supplement our evidence development
  - *Hypothesis generating, confirmatory, extension of trials to understudied subsets*
  - *Must be careful consumers → some treatment comparisons may not be possible in observational data (at least with traditional methods to adjust for confounding)*

Designing Clinical Research

- Randomized trials vs. observational studies
- Superiority vs. Non-inferiority trials
- Blinding and placebos

Terminology

Superiority Trial

– Prototypical clinical trial where the goal is to demonstrate that the new treatment is better than placebo or standard therapy

Non-Inferiority Trial

– Trial design where the goal is to show that the new therapy is not worse than standard therapy by some tolerable margin (e.g., 30-day mortality difference no greater than 1%)

Why perform a non-inferiority trial?

- Placebo control trial unethical but still want to demonstrate that the new treatment is better than nothing ("putative placebo") approach
- New therapy may offer important advantages over currently available effective therapies
  - *Improved safety*
  - *Better tolerability/fewer side effects*
  - *Ease of use (2nd generation DES, QD drug, etc.)*
  - *Less expensive*
  - *Increased market competition (?)*

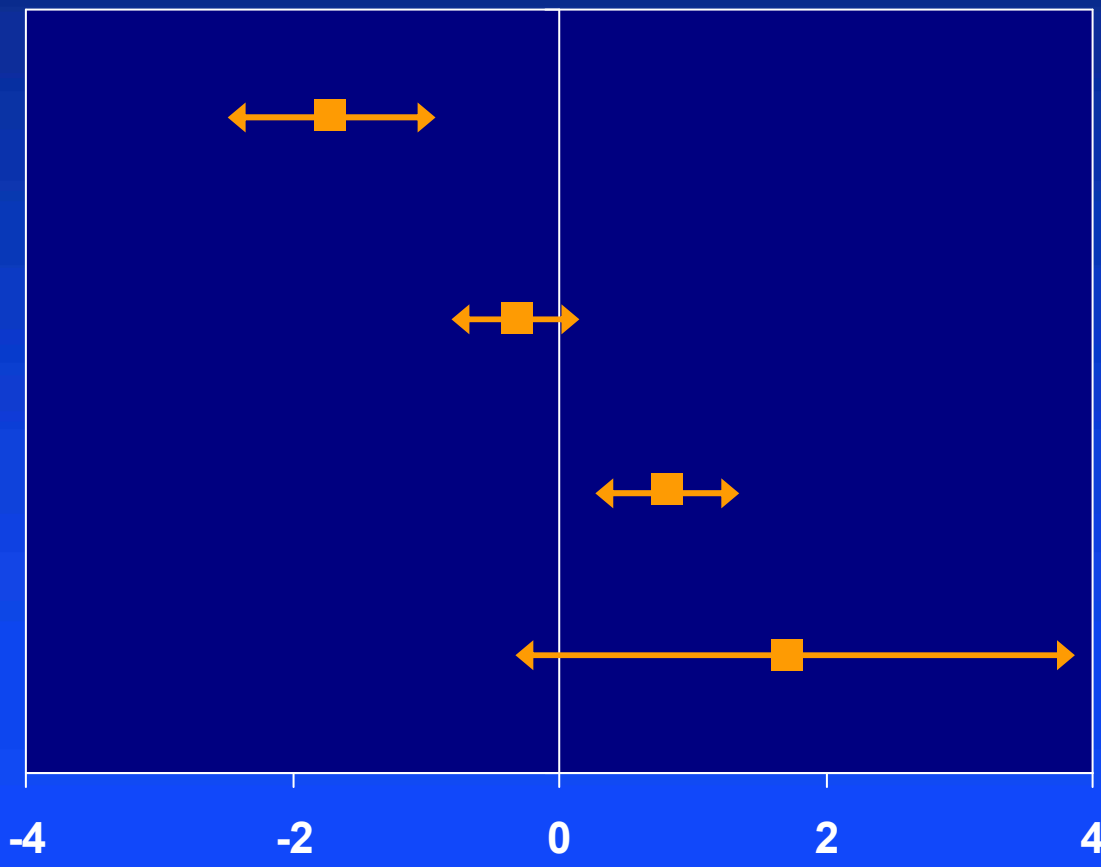
How can you prove equivalence?

Statistical Testing: Superiority Trial

H_0 (Null Hypothesis)	H_a (Alternate Hypothesis)
$E_T = E_S$	$E_T \neq E_S$

Application: If we can reject the null hypothesis (with 95% certainty), this represents strong evidence that the 2 treatments are not equivalent (and that one or the other is superior)

Statistical Concepts: Superiority Trial



Interpretation
Inferior
Uncertain
Superior
Uncertain

Statistical Testing: Non-Inferiority Trial

H_0 (Null Hypothesis)

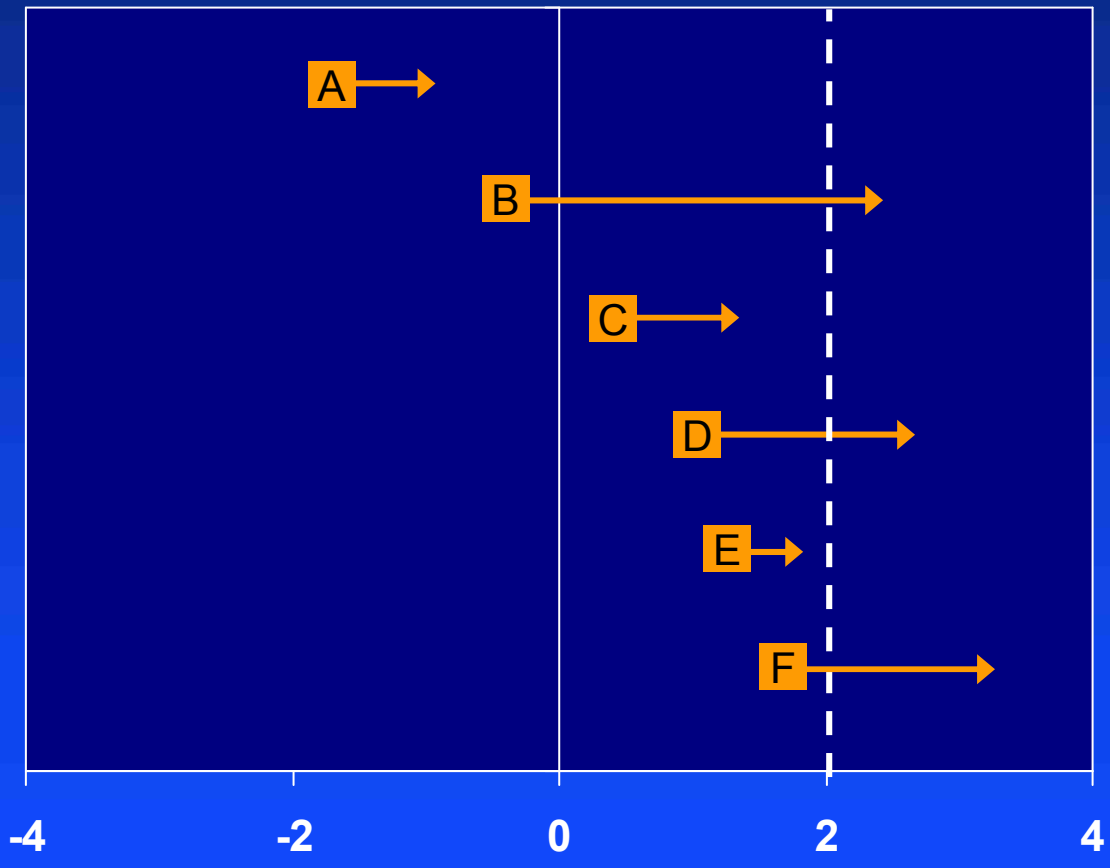
$E_T - E_S \geq \delta$

H_a (Alternate Hypothesis)

$E_T - E_S < \delta$

Application: If we can reject the null hypothesis (with 95% certainty), this provides strong evidence that the test treatment is not worse than the standard treatment by δ (the non-inferiority margin)

Statistical Concepts: Non-Inferiority Trial



Interpretation

A	Non-inferior
B	Not non-inferior
C	Non-inferior
D	Not non-inferior
E	Non-inferior
F	Not non-inferior

1-sided test
Alpha = 2.5%

Selecting a non-inferiority margin

- Critical to pre-specify the non-inferiority margin to avoid Type I error (false positive results)
- Potential approaches
  - Clinical rationale→ expert opinion ("what is the maximum difference you would tolerate?")
  - Regulatory rationale→ based on previous trials
  - Statistical + Clinical rationale→ designed to preserve some minimum proportion of benefit vs. placebo ("putative placebo" approach)
- Rule of thumb: Margin cannot be greater than the smallest effect size that the active comparator would be expected to have vs. placebo

Designing Clinical Research

- Randomized trials vs. observational studies
- Superiority vs. Non-inferiority trials
- Blinding and placebos

Randomization alone may not be enough

- Randomization is a technique for eliminating confounding (both measured and unmeasured)
- However, randomization does not eliminate bias

What is bias?

- Definition: Any systematic difference in the way that subjects in a study are handled
- Examples:
  - *Selection bias → patients with certain characteristics not entered into trial*
  - *Treatment bias → patients treated differently in 2 arms of the trial (e.g., different medications, different f/u, etc.)*
  - *Ascertainment bias → outcomes are assessed differently depending on the treatment assignment*

Blinding is a technique for eliminating bias

Who to blind

- Person enrolling patient → if they know the “next” treatment assignment, they may try to select a specific type of patient who would be expected to respond well to that treatment
- Patient
- Personnel involved in follow-up care
- Personnel involved in assessing study endpoints (e.g., angiographic core laboratory, clinical events committee)

When and how to blind

- Blinding assumes increasing importance with the degree of subjectivity of the endpoint
- Examples:
  - *All-cause mortality: Little potential for bias*
  - *Angiographic restenosis: reasonably objective, still need to blind the core laboratory*
  - *Repeat Revascularization: strong potential for bias → blinding of patient and physician/assessor critical*
- Use of placebo (or sham procedures) is the optimal method to maintain blinding
  - *Not always feasible, however, if the treatment is highly invasive or the medication has characteristic side effects*

Summary

- Like clinical medicine, clinical research is both an art and a science
- No single study design will suit all possible questions
- Key factors to consider in every study:
  - *What is the appropriate comparator and type of comparison?*
  - *Size of expected treatment effect– small to moderate effects will require randomization to minimize confounding*
  - *Is blinding needed→ more important with subjective endpoints*
- The role of the clinical investigator is to integrate all of these factors to develop a practical, feasible, and cost-effective study design