Main Line Health®
Well ahead.®

How to set-up randomized trials:

Superiority vs. Non-inferiority Designs

William A. Gray MD
System Chief of Cardiovascular Services, Main Line Health
President, Lankenau Heart Institute
Wynnewood, PA
USA

The premise (and promise) of randomized controlled trials (RCTs)

- Proper randomization ***minimizes allocation bias***, balancing both known and unknown prognostic factors, in the assignment of treatments being studied
  - eliminates bias in treatment assignment, specifically selection bias and confounding
  - facilitates blinding (masking) of the identity of treatments from investigators, participants, and assessors
  - ***permits the use of probability theory to express the likelihood that any difference in outcome between treatment groups merely indicates chance***

The “C” stands for “control”

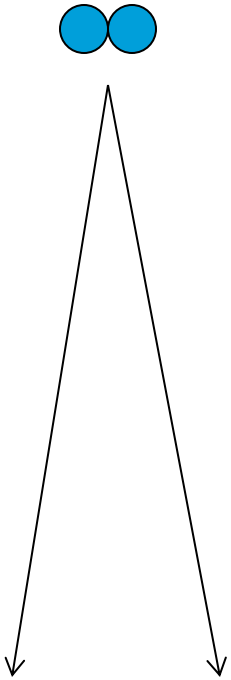
- The control group may receive a placebo treatment, but more commonly in device trials the control group receives an established, previously tested, treatment
- The ethics of RCTs dictate that:
  - the control treatment reasonably represent the standard of care
  - a reasonable uncertainty exists among the investigators that the treatments being offered are roughly equivalent in safety and efficacy (clinical equipoise)

Classifications of RCTs: by study design

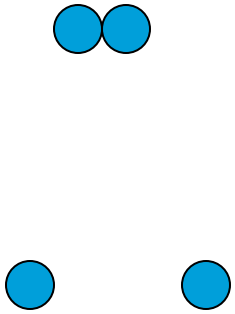
- **Parallel group:** each participant is randomly assigned to a treatment group
- **Cross-over:** subjects receive a sequence of different treatments
- **Cluster:** pre-existing groups of participants (e.g., villages, schools) are randomly selected to receive (or not receive) an intervention
- **Factorial:** each participant is randomly assigned to a group that receives a particular combination of interventions or non-interventions

RCT classifications

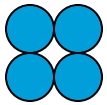
Parallel



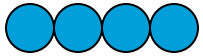
Cross-over



Cluster



Factorial



Classifications of RCTs: by outcome of interest

- RCTs can be classified as "explanatory" or "pragmatic."
  - Explanatory RCTs test *efficacy* in a research setting with highly selected participants and under highly controlled conditions (IMPERIAL).
  - Pragmatic RCTs test *effectiveness* in everyday practice with relatively unselected participants and under flexible conditions; in this way, pragmatic RCTs can inform decisions about practice (ACST 2)

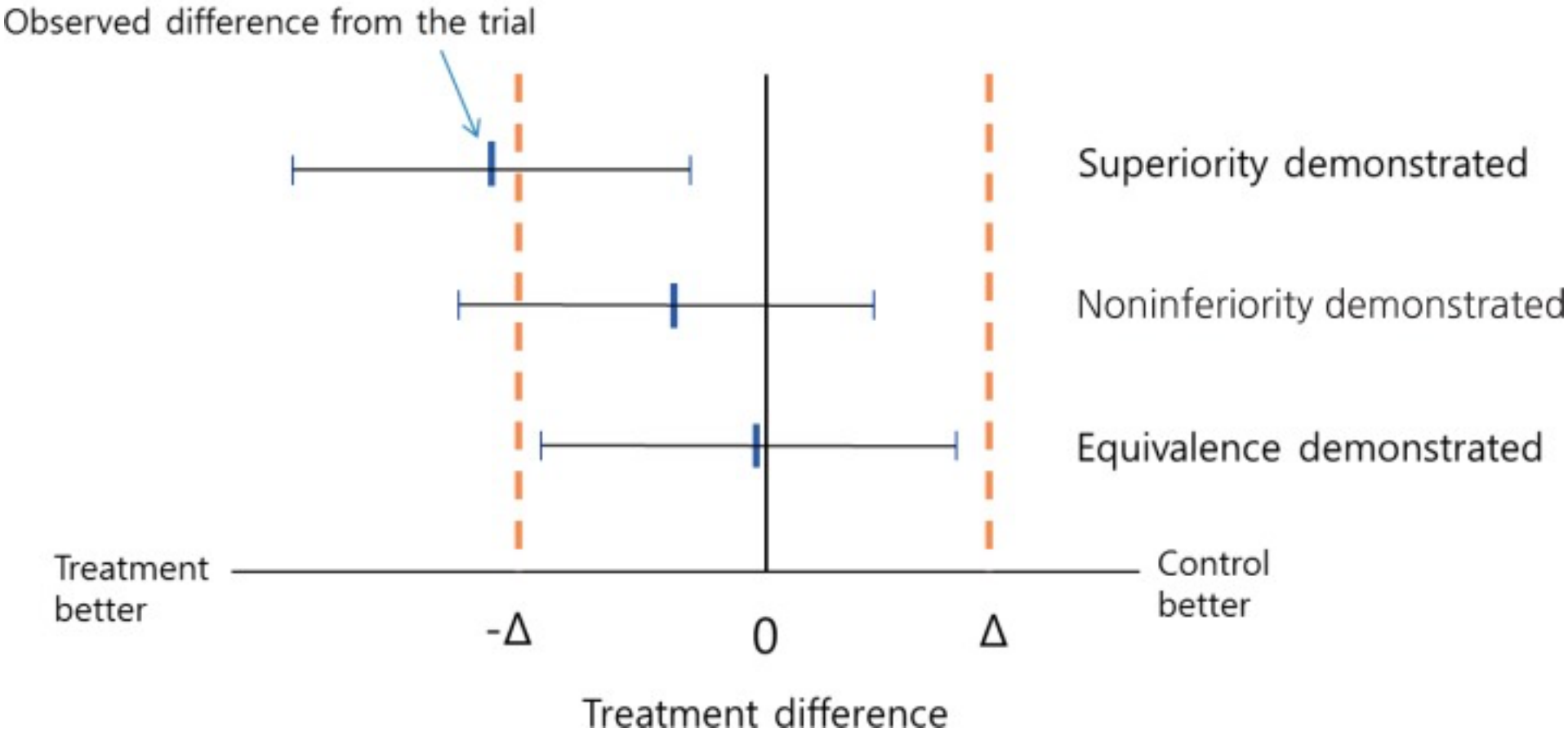
Classifications of RCTs: by hypothesis

- RCT hypothesis categories include superiority trials, noninferiority trials, and equivalence trials---these differ in methodology and reporting.
  - Most RCTs are ***superiority*** trials, in which one intervention is hypothesized to be superior to another in a statistically significant way.
  - Some RCTs are ***noninferiority*** trials to determine whether a new treatment is no worse than a reference treatment.
  - Even fewer RCTs are ***equivalence*** trials in which the hypothesis is that two interventions are indistinguishable from each other

Some statistical mumbo-jumbo

- Superiority trials:
  - Seek sufficient evidence to reject the hypothesis that "the 2 treatments have equal effects" in favor of the superiority of the new treatment
  - Failure to demonstrate superiority does not necessarily translate into non-inferiority or equivalence
- Non-inferiority trials:
  - Constructed to reject the null hypothesis that the experimental treatment is inferior by an equivalence margin (i.e., 95% CI limit)

95% Confidence interval noninferiority



Classification of RCTs: US-based commentary

- FDA will typically drive much of the decision-making around trial design and conduct
- CMS is increasingly unwilling to reimburse for technology which more expensive but not clearly superior to existing therapies
  - Unless it can be amply demonstrated to have better safety, convenience, or compliance

Superiority trials and ITT analysis (Pros)

- In ***superiority trials***, intention-to-treat (ITT) is the standard unbiased analyses:
  - ITT really reflects a trial strategy that involves the design, conduct and analysis---not just the analysis
  - ITT analyzes all patients according to their assigned treatment, regardless of whether they received them, thus:
    - Represents a “practical’ clinical situation
    - Does not become confounded by unbalanced missing data
    - Guards against a Type I error (false positive)
  - ITT preserves the sample size and thus the statistical power of the study

Superiority and ITT analysis

- Presumes no/minimal lost to follow-up or crossover, only then can true “pure” ITT analysis be performed
  - In the case of excessive lost to follow-up and/or crossover, modified ITT (mITT) can be used, but can be arbitrary in selection of patients to be excluded
  - “Last observation carried forward” (LOCF) and multiple imputation methods can also be used to account for missing data
- ITT the most conservative approach to proof of superiority
  - But in cases of lost data, may be prone to Type II error (false negative results)

Non-inferiority trials: different analyses may be required

- In non-inferiority trials, per protocol (PP) analyses are often used
  - This represents the most conservative assessment of non-inferiority
    - If lost subjects were included, then it would skew results toward non-inferiority
  - PP includes only subjects that fulfill the protocol in the terms of the eligibility, interventions, and outcome assessment

Practical example in modern endovascular trials: CREST

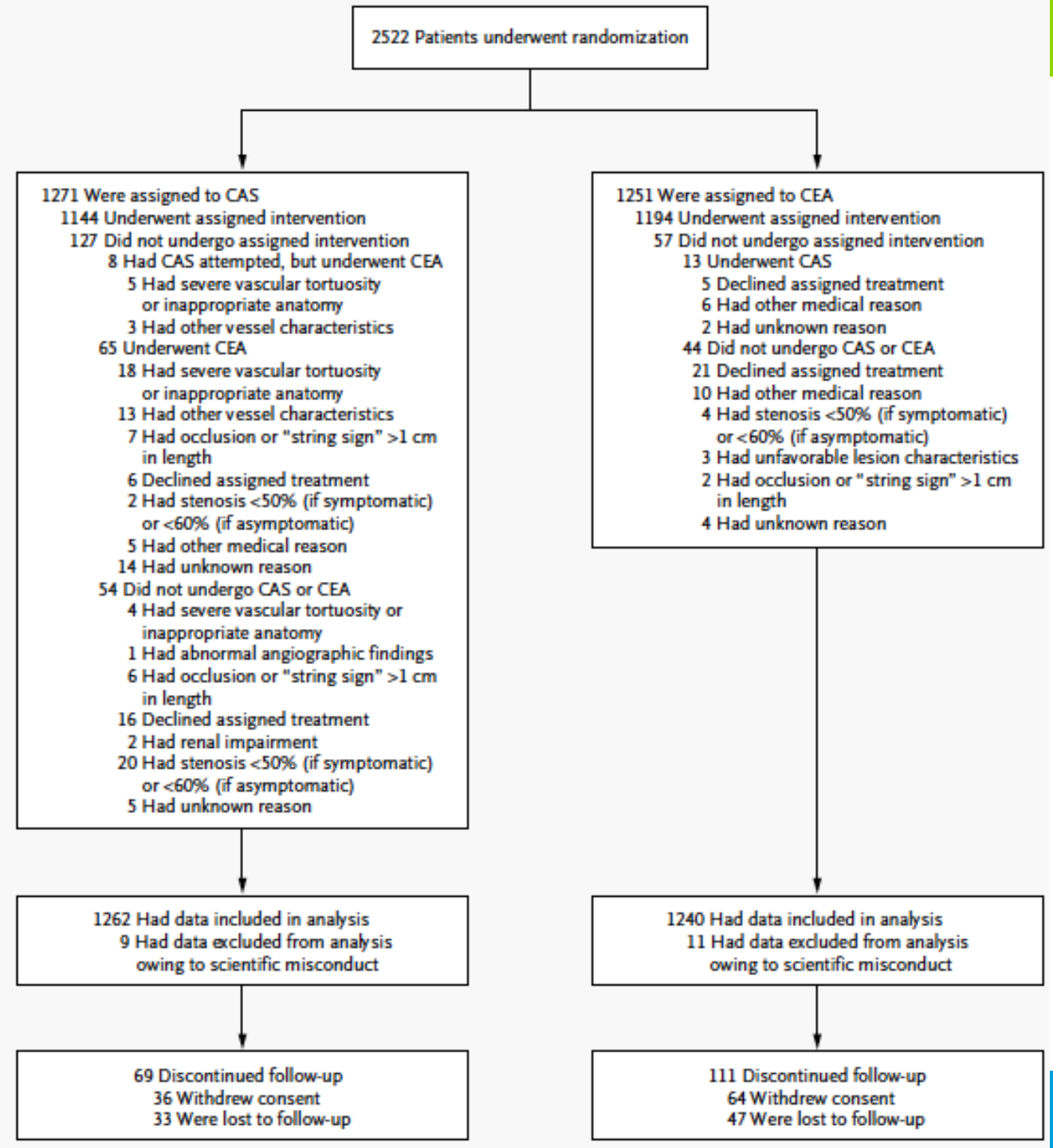
- NIH analysis: superiority, ITT analysis

Brott et al. Stenting vs. endarterectomy for the treatment of carotid artery stenosis. N Eng J Med. 2010; 363(1): 11-23

- PMA (FDA) analysis: non-inferiority, PP analysis

Gray et al. Overview of the 2011 FDA Circulatory System Devices Panel Meeting on the ACCULINK and ACCUNET Carotid Artery Stent System. Circulation. 2012: 125: 2256-2264

CONSORT chart documenting patient flow



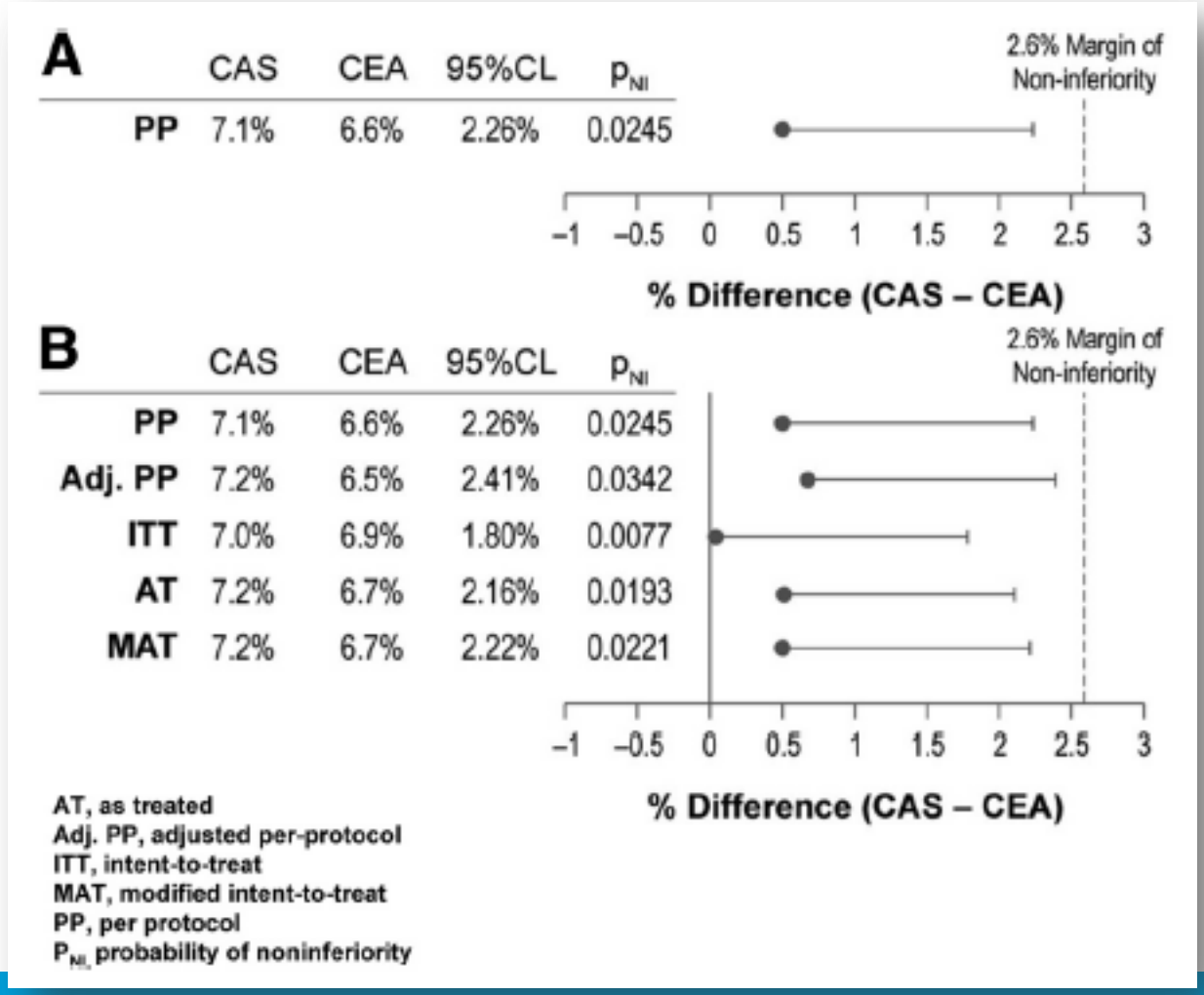
Differences in event rates based on PP vs. ITT analysis

Table 3. Primary Endpoint Components for NIH and PMA Analyses

	PMA Analysis (within 30-day)		NIH Analysis (Peri-procedural)*	
	CAS N = 1131	CEA N = 1176	CAS N = 1262	CEA N = 1240
Death, Stroke or MI	5.8% (65)	5.1%(60)	5.2% (66)	4.5% (56)
Any Stroke	4.1% (46)	1.9% (22)	4.1% (52)	2.3% (29)
Major Stroke	0.9% (10)	0.4% (5)	0.9% (11)	0.6% (8)
Minor Stroke	3.2% (36)	1.5% (18)	3.2% (41)	1.7% (21)
MI	2.0% (22)	3.4% (40)	1.1% (14)	2.3% (28)
Death	0.53% (6)	0.26% (3)	0.7% (9)	0.3% (4)

Practical example in modern endovascular trials: CREST results

Non-inferiority demonstrated in multiple analyses



Summary

- The choice of trial design—superiority vs. non-inferiority--- will be based on:
  - Intrinsic factors: device/disease/available comparators
  - Extrinsic factors: regulatory and reimbursement requirements/environments
- The type of trial will dictate not only the statistical analysis (e.g., ITT vs. PP) but also the philosophy of trial design and conduct
- Our interpretation of the results of any trial should keep these facts in mind

Main Line Health®
Well ahead.®

How to set-up randomized trials:

Superiority vs. Non-inferiority Designs

William A. Gray MD
System Chief of Cardiovascular Services, Main Line Health
President, Lankenau Heart Institute
Wynnewood, PA
USA