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

- Superiority demonstrated
- Noninferiority demonstrated
- Equivalence demonstrated

Treatment difference

Treatment better

Control better
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


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

CONSORT chart documenting patient flow
Table 3. Primary Endpoint Components for NIH and PMA Analyses

<table>
<thead>
<tr>
<th></th>
<th>PMA Analysis</th>
<th>NIH Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(within 30-day)</td>
<td>(Peri-procedural)*</td>
</tr>
<tr>
<td></td>
<td>CAS</td>
<td>CEA</td>
</tr>
<tr>
<td>N</td>
<td>N = 1131</td>
<td>N = 1176</td>
</tr>
<tr>
<td>Death, Stroke or MI</td>
<td>5.8% (65)</td>
<td>5.1% (60)</td>
</tr>
<tr>
<td>Any Stroke</td>
<td>4.1% (46)</td>
<td>1.9% (22)</td>
</tr>
<tr>
<td>Major Stroke</td>
<td>0.9% (10)</td>
<td>0.4% (5)</td>
</tr>
<tr>
<td>Minor Stroke</td>
<td>3.2% (36)</td>
<td>1.5% (18)</td>
</tr>
<tr>
<td>MI</td>
<td>2.0% (22)</td>
<td>3.4% (40)</td>
</tr>
<tr>
<td>Death</td>
<td>0.53% (6)</td>
<td>0.26% (3)</td>
</tr>
</tbody>
</table>

Differences in event rates based on PP vs. ITT analysis

Lankenau Heart Institute
Main Line Health
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
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