Clinical Trial Design

Statistical Approaches and Considerations

Michael J. Pencina, PhD

Duke Clinical Research Institute

Disclosures

- Industry grants to Duke Clinical Research Institute
- Consulting fees from Pamlab Inc

Outline

- Numerous statistical and design approaches exist to increase trial efficiency
- Many have been tried in cardiology, oncology and other areas
- We review a selection here including
  - Event driven trials
  - Composite outcome
  - Adaptive designs
  - Enrichment designs
  - Borrowing of controls
  - Opportunities in era of big data

Event driven trials

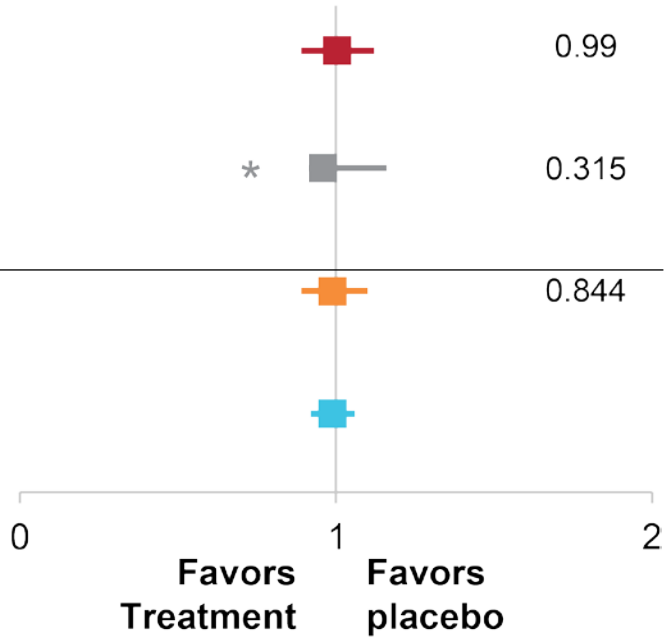
- In time-to-event outcome studies power depends on the number of events observed
- Efficiency is increased by stopping when minimum necessary number of events is reached
- All follow-up is included in the analysis
- Blinded interim monitoring of event count allows increasing or decreasing planned study duration or increase in enrolled sample

Composite Outcomes

- Number of primary outcome events can be increased by combining different outcome types
- For example, major adverse cardiovascular events (MACE) consist of cardiovascular death, myocardial infarction and stroke
- Time to first event or hierarchical composite used
- They should be of similar severity
- Analysis of individual components always conducted as sensitivity

Example: Time to MACE composite

	Study Drug n/N (%)	Placebo n/N (%)	Hazard Ratio	95% CI	P Value
SAVOR-TIMI (saxagliptin vs. placebo)	613/8280 (7.4%)	609/8212 (7.4%)	1.00	0.89, 1.12	0.99
EXAMINE (alogliptin vs. placebo)	305/2701 (11.3%)	316/2679 (11.8%)	0.96	NA, 1.16	0.315
TECOS (sitagliptin vs. placebo)	745/7332 (10.2%)	746/7339 (10.2%)	0.99	0.89, 1.10	0.844
SAVOR + EXAMINE + TECOS	1663/18313 (9.1%)	1671/18230 (9.2%)	0.99	0.92, 1.06	



1. Scirica BM et al. N Engl J Med 2013; 369: 1317–1326
2. White WB et al. N Engl J Med 2013; 369: 1327–1335
3. Green JB et al. NEJM 2015; DOI: 10.1056/NEJMoa1501352

Example: TAVI in aortic stenosis

- The Placement of Aortic Trans-catheter Valves (PARTNER) trial used Finkelstein-Schoenfeld hierarchical composite
- All-cause mortality first in hierarchy, hospitalization for heart failure second
- Multiple pairwise comparisons performed for all patient pairs, first with respect to time to death and then with respect to time to repeat hospitalization
- Method more powerful if some outcomes continuous

1. Leon MB *et al*. N Engl J Med 2010; 363:1597-1607
2. Finkelstein DM, Schoenfeld DA. Stat Med 1999;18:1341-1354

Adaptive Designs

- Key study features (sample size, study duration, number of treatment arms) can be adapted based on information obtained at interim
- Blinded interim looks focus on combined parameters (event rate, overall mean, variance) and usually do not trigger alpha penalty
- Unblinded interim looks incorporate effect size observed at interim and usually trigger alpha penalty
- Decision often based on conditional power

Example: drop-the-losers design

- In stage A, k experimental and one control treatment administered
- Data unblinded and analyzed and only the best treatment or any treatment exceeding pre-specified threshold proceed to stage B along with control
- Final summary statistic based on sample size-weighted combination of effect versus placebo from stages A and B

Sampson and Sill, Biometrical Journal 2005 47(3): 257–268

Biomarkers

- FDA's Qualification process for drug development tools:
  - A *prognostic* biomarker is a baseline patient or disease characteristic that categorizes patients by degree of risk for disease occurrence or progression. A prognostic biomarker informs about the natural history of the disorder in that particular patient in the absence of a therapeutic intervention
  - A *predictive* biomarker is a baseline characteristic that categorizes patients by their likelihood for response to a particular treatment. A predictive biomarker is used to identify whether a given patient is likely to respond to a treatment intervention in a particular way (favorable or not)

Example: Predictive biomarkers

Effect of l-Methylfolate 15 mg/d vs Placebo on Pooled Mean Change From Baseline for HDRS-28 Stratified by Baseline Level of Plasma Marker				
Marker	N	Pooled Mean Change vs. Placebo	95% CI	P-value
SAM/SAH ratio				
≥ 2.71 mg/L	36	0.07	-3.33, 3.48	0.966
< 2.71 mg/L	37	-4.57	-7.73, -1.41	0.005
4-HNE				
≥ 3.28 µg/mL	37	-4.55	-7.61, -1.50	0.003
< 3.28 µg/mL	36	-0.11	-3.67, 3.46	0.953

Papakostas et al. J Clin Psychiatry 2014; 75:8

Adaptive Enrichment Designs

- G_0 – all eligible subjects
- G_1 – subset of all eligible subjects who possess non-wild-type alleles on either genotype A and B
- G_2 – subset of eligible subjects who possess non-wild-type alleles on both genotype A and B
- Conditional power calculated at interim based on effect size in each set
- Opportunity to enrich by recruiting only to subset G_1 or G_2

Sampson and Sill, Biometrical Journal 2005 47(3): 257–268

Sequential Parallel Comparison Design

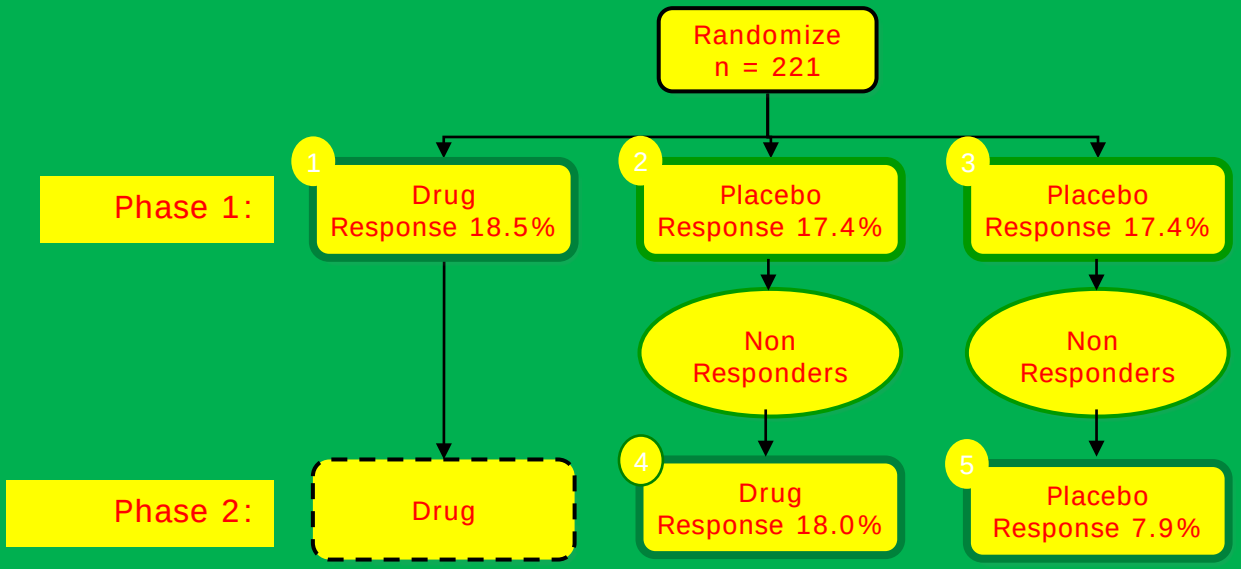
- Developed to reduce impact of placebo response
- In stage A, individuals randomized to treatment versus placebo (usually more to placebo)
- Stage A placebo non-responders re-randomized to treatment versus placebo
- Final summary statistic based on weighted combination of effects from stages A and B
- Particular choice of weights enables interpretation as effect after accounting for placebo response

Fava M et al. Psychother. Psychosom. 72 (2003): 115-227.
Doros G et al. Statistics in Medicine 32 (2013): 2767–2789.
Chi G. Contemporary Clinical Trials Communications 2 (2016): 34-53

ADAPT-A: Aripiprazole Augmentation of SSRIs

Source: Fava et al, Psychotherapy and Psychosomatics 2012

Response rate	Drug	Placebo	Difference
SPCD Phase 1	18.5%	17.4%	1.1%
SPCD Phase 2	18.0%	7.9%	10.1%



Borrowing Controls

- All or some controls are borrowed from “historical” data
- Numerous options:
  - Pooling: adds historical controls to randomized controls
  - Performance criterion: uses historical data to define performance criterion for current, treated-only trial to beat
  - Test then pool: test if controls sufficiently similar for pooling
  - Power priors: historical control discounted when added to randomized controls
  - Hierarchical modeling: variation between current vs. historical data is modeled in Bayesian fashion

Opportunities in era of big data

- We might be able to run very large simple trials for fraction of cost
  - EHR-enabled trials
    - pcornet ADAPTABLE – trial of aspirin doing in secondary prevention with 20,000 patients with EHR as primary source of data capture
  - Registry-based trials
    - SAFE-PCI in women – comparing access site (radial vs. femoral) in 1800 PCI or angiography with possible PCI women
    - Several ongoing Registry-RCTs in Sweden

ADAPTABLE Study Design

Patients with known ASCVD + ≥1 "Enrichment Factor"

Identified through EHR screening and electronic patient contact by CDRNs/PPRNs
(PPRN patients would need to connect through a CDRN to participate)

Patients contacted electronically with trial information and e-consent via web portal
Treatment assignment will be provided directly to patient

***Enrichment Factors**

- Age > 65 years
- Creatinine > 1.5 mg/dL
- Diabetes mellitus (type 1 or 2)
- Known 3-vessel CAD
- Current CVD or PAD
- Known EF<50% by echo, cath, nuclear study
- Current smoker

ASA 81 mg QD

ASA 325 mg QD

Randomized Electronic Follow-Up: 3 vs 6 months
Supplemented with EHR/CDM Data Queries

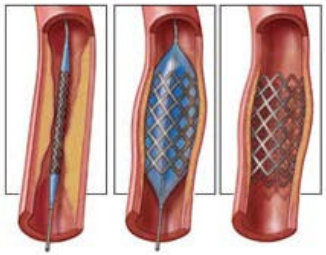
Duration: Enrollment over 24 months; maximum follow up of 30 months

Primary Endpoint: Composite of all-cause mortality, hospitalization for MI, or hospitalization for stroke

Primary Safety Endpoint: Hospitalization for major bleeding

Methods - SAFE-PCI for Women Workflow

Randomization



Demographics
Medical History
Procedural data

NCDR® National Cardiovascular Data Registry
CathPCI Registry®

Autopopulate trial database with registry "big" data

Unique pages for trial

ORACLE

Analytic Database

DCRI

Conclusions

- Numerous study designs and statistical approaches intend to approve trial efficiency
- Different approaches at different stages of adoption and different potential for application in neuroscience studies
- In general, regulators more open to innovative approaches in smaller studies and/or earlier stages of development (II vs. III)
- Careful consideration and appropriate statistical expertise in the planning stage needed to identify approaches most likely to succeed