Statistics for Clinical Trials: Basics of a Phase III Trial Design

Greg Pond

Ph.D., P.Stat.

Ontario Clinical Oncology Group
Escarpment Cancer Research Institute
Department of Oncology, McMaster University
Level II Investigator, Ontario Institute for Cancer Research

07 August 2019

McMaster University
HEALTH SCIENCES

Conflicts

- Family member works for Roche Canada and owns stock
- Received honorarium (DSMB) from Takeda

McMaster University
HEALTH SCIENCES

Phase III RCTs

1. Two-arm, double-blind, superiority RCT
2. Cross-over
3. Biomarker-enriched
4. Biomarker stratified
5. Equivalence / Non-inferiority
6. Factorial
7. Bayesian adaptive
8. Dynamically allocated randomization
9. …

McMaster University
HEALTH SCIENCES

Outline

1. Randomization
2. Blinding / Concealment
3. Intention-to-treat
4. Statistical Power (α and β)
5. P-values / Confidence Intervals

McMaster University
HEALTH SCIENCES

McMaster University
HEALTH SCIENCES

What is Randomization?

- The most fundamental principle in statistics
- Ensures comparability of interventions
- Non-deterministic process by which patients assigned to intervention
- All patients have same chance of getting each treatment

McMaster University
HEALTH SCIENCES

Statistical Importance

- Patient cohorts become similar / balanced as sample size get large
- Balanced: known and *unknown* characteristics
- Observed difference (outcomes) due to treatment effect / imbalance in characteristics / chance
- Chance can be quantified

McMaster University
HEALTH SCIENCES

Statistical Interpretation

- Median survival for patients given A > median survival for patients given B by 6 months

- Prob(due to chance) = XX (say 0.002)

- Prob(baseline imbalance) = Prob(due to chance)

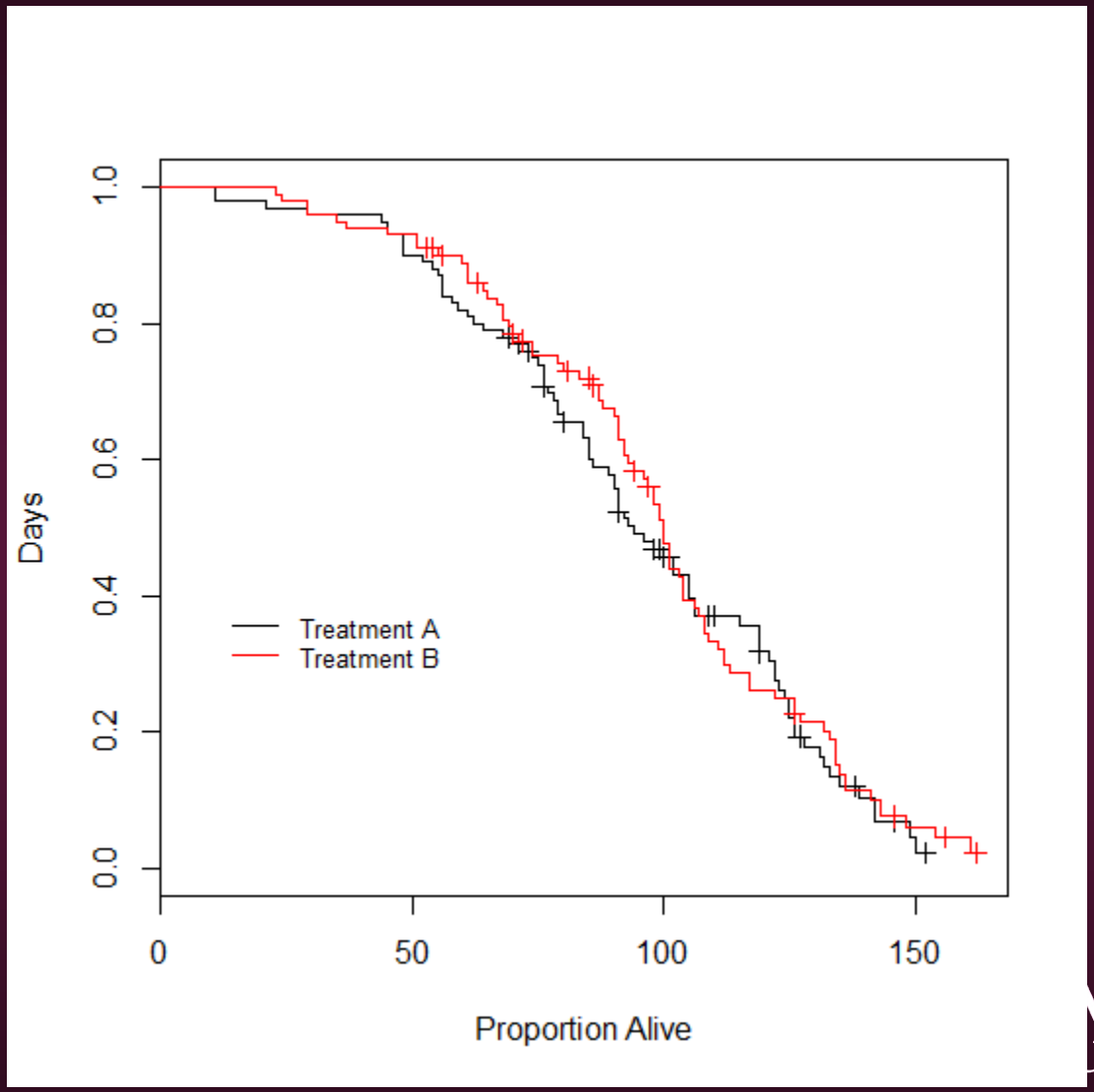
- Prob(treatment effect)=?, however, chance is unlikely => assume treatment effect

McMaster University
HEALTH SCIENCES

Problems

- True randomization balances as sample sizes get large
- Many clinical trials have small sample sizes
- Unequal # of pts allocated to each arm (cost, feasibility)
- Imbalance in characteristics (credibility)

McMaster University
HEALTH SCIENCES



McMaster University
HEALTH SCIENCES

Quasi-Randomization

McMaster University
HEALTH SCIENCES

Quasi-Randomization

- Permuted Blocks Random Sampling
- AABB, ABBA, ABAB, BBAA, BAAB, BABA
- Randomly select a block
- Ensures approximately equal numbers of patients get each treatment

McMaster University
HEALTH SCIENCES

Quasi-Randomization

- Stratified Random Sampling
- Select ‘stratification factors’ of importance
- Permuted blocks within strata
- Ensures approximately even number of patients within each stratum receive each treatment

McMaster University
HEALTH SCIENCES

Dynamic Allocation

- Often referred to as minimization
- Evaluate characteristics of patients already on study
- Allocate next patient to treatment which will create better balance
- e.g. if 10 women received A and 7 received B, then the next woman allocated to B

McMaster University HEALTH SCIENCES

Blinding / Concealment

- Important that researchers do not know next allocation
- If I know permuted blocks used, and last three patients were AAB, then I know next patient is B
- I may (sub)consciously (not) recommend trial to next patient
- Bias trial results

McMaster University
HEALTH SCIENCES

Subconscious Example

- Organize clinic so easier cases earlier in the day
- Personal belief that A is easier to tolerate (despite community equipoise)
- Know next patient will get B
- May be less likely to present trial to complex patient at end of day (7 PM)

McMaster University HEALTH SCIENCES

Terminology

- Concealment – ensuring people are unaware of next treatment allocation

- Blinding – ensuring people are unaware of treatment patient is receiving

- Reduces ability to 'guess' next treatment

- Reduces bias caused by process changes (e.g. schedule changes / conmeds)

McMaster University
HEALTH SCIENCES

Blinding Terminology

- Single blind – patients do not know treatment (surgery trials)
- Double blind – patients and physicians blinded
- Triple blind – adjudicators blinded also (radiologist); May change treatment incorrectly
- Reduces bias, but less similar to real life

McMaster University
HEALTH SCIENCES

Blinding considerations

- Must consider ability to blind
- Surgery vs oral medication
- Will toxicities unblind allocation?
- How does blinding affect future treatments?

McMaster University
HEALTH SCIENCES

Intent-to-Treat Principle

- All patients randomized should be analyzed according to *allocation*

- Patient randomized to surgery (experimental arm). Opts to withdraw and gets oral medication (control arm)

- Analyzed on surgery arm

McMaster University
HEALTH SCIENCES

Intent-to-Treat Principle

- More conservative analysis. E.g. assume all patients cross over and get same treatment. Treatment effect is 0.

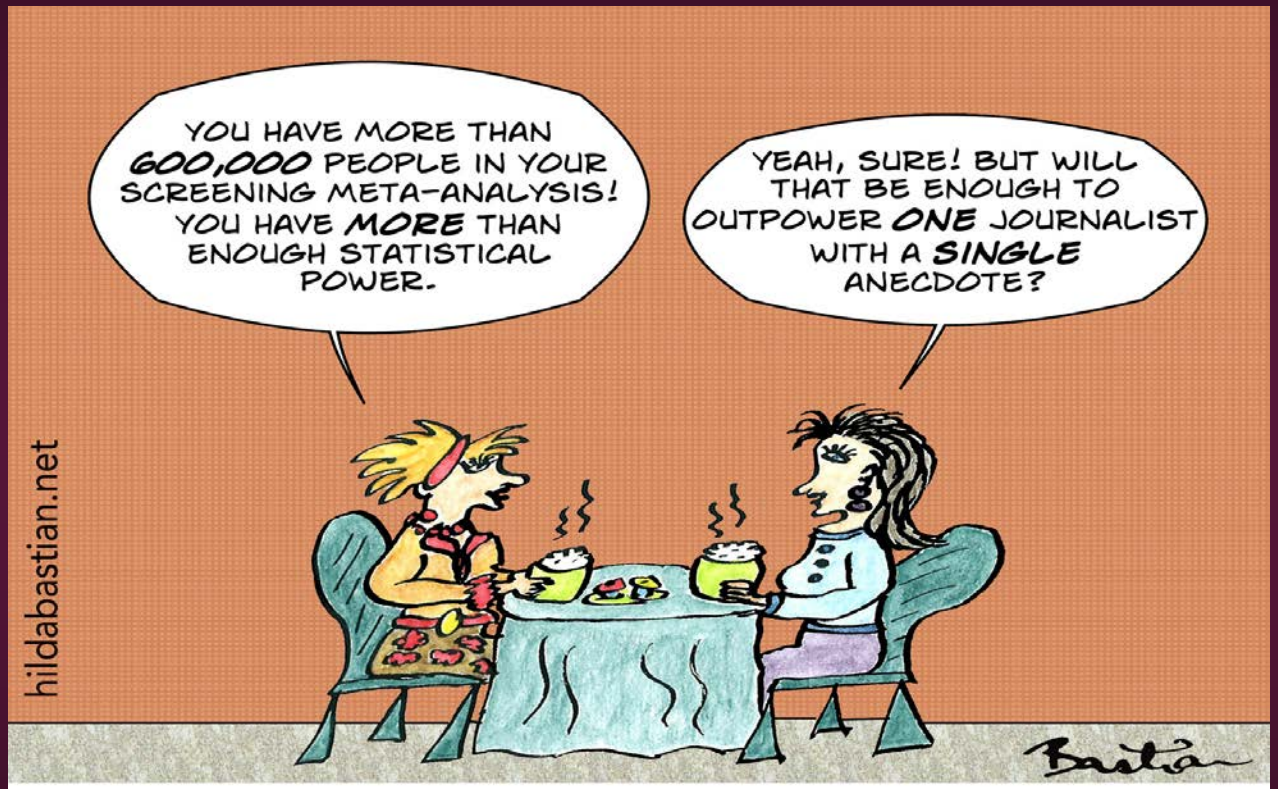
- ITT biases towards no difference. Hence, if H0 rejected, we have strong evidence to do so.

- Reduces bias due to perceived / true lack of blinding

- Preserves planned statistical power

McMaster University
HEALTH SCIENCES

Statistical Power



McMaster University
HEALTH SCIENCES

Error Rates

- α is the probability, *assuming H0 is true*, that we will reject H0
- *If drug is inactive*, α is the probability our study will conclude the drug is active
- β is the probability, *assuming HA is true*, that we will not reject H0
- *If drug is active*, β is the probability our study will conclude the drug is inactive
- Power is (1-β)*100%

McMaster University
HEALTH SCIENCES

Statistical Errors

	Truth is H0	Truth is HA
Study Conclusion is H0 is True	Study is Correct	β, type II
Study Conclusion is HA is True	α, type I	Study is Correct

McMaster University
HEALTH SCIENCES

Statistical Errors

	Truth is H0	Truth is HA
Study Conclusion is H0 is True	Study is Correct	β, type II
Study Conclusion is HA is True	α, type I	Study is Correct

At study conclusion, only information available

McMaster University
HEALTH SCIENCES

Designing a Study

- Want to increase chance of making a correct decision
- For fixed sample size, if α decreases, β increases
- To decrease α and β, must increase sample size
- Sample size calculation is the minimum number required so that both α and β are ≤ some 'reasonable value'

McMaster University
HEALTH SCIENCES

Why α=0.05, β=0.20?

- No statistical motive, but 'works'
- α error: Truth is novel agent is inactive

=>Further study in phase III, patient/financial costs

- β error: Truth is novel agent is active

=>Not studied again, lost a potentially useful treatment

- Weigh relative cost of each error

McMaster University
HEALTH SCIENCES

P-values

- P-values: Probability, assuming H0 is true, of observing data as extreme or more extreme, than what actually was observed if trial was repeated identically many times

- NOTE: there is an ongoing debate amongst statisticians whether p-values should be reported or not!

Problems with p-values

- Does not say it is true, just it is plausible
- 'we do not have enough evidence to reject H0"

- Low p-values do NOT mean H0 is false
- Assumption: probability is low

=>unlikely to occur by chance

=>more likely that H0 is false

McMaster University
HEALTH SCIENCES

Problems with p-values

- P-values of 0.051 is not really different than p-values of 0.049

- Except p=0.049 gets a better publication

McMaster University
HEALTH SCIENCES

P-values are probabilities

- Probabilities have different meanings, depending on the context

- Assume patient has positive test result from a diagnostic test, false-positive rate=0.01

- Then take a second test which is negative, and false-negative rate=0.0000001

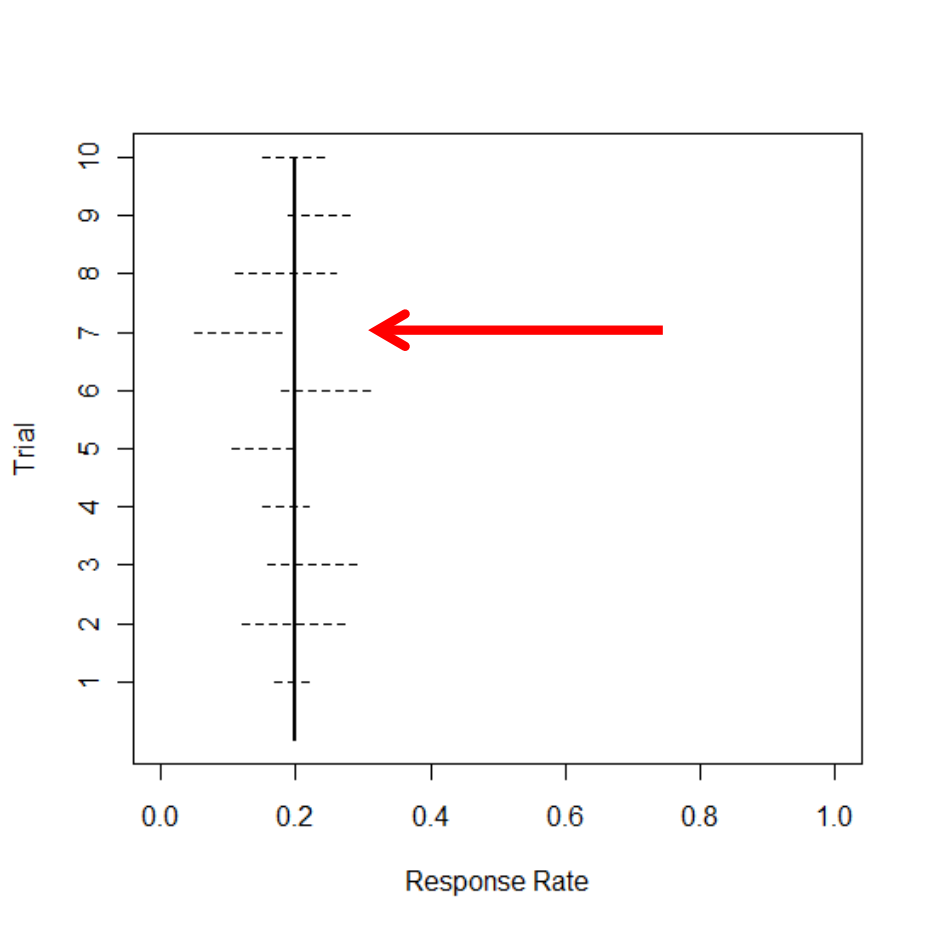
McMaster University
HEALTH SCIENCES

Confidence Intervals

- Range of values of which the data are consistent
- If H0 is any value within a 95% CI, then the p-value would be ≥0.05
- It does NOT mean the true value is in CI
- If trial repeated many times – 95% of identically constructed CI will cover true value

McMaster University
HEALTH SCIENCES

Confidence Intervals



McMaster University
HEALTH SCIENCES

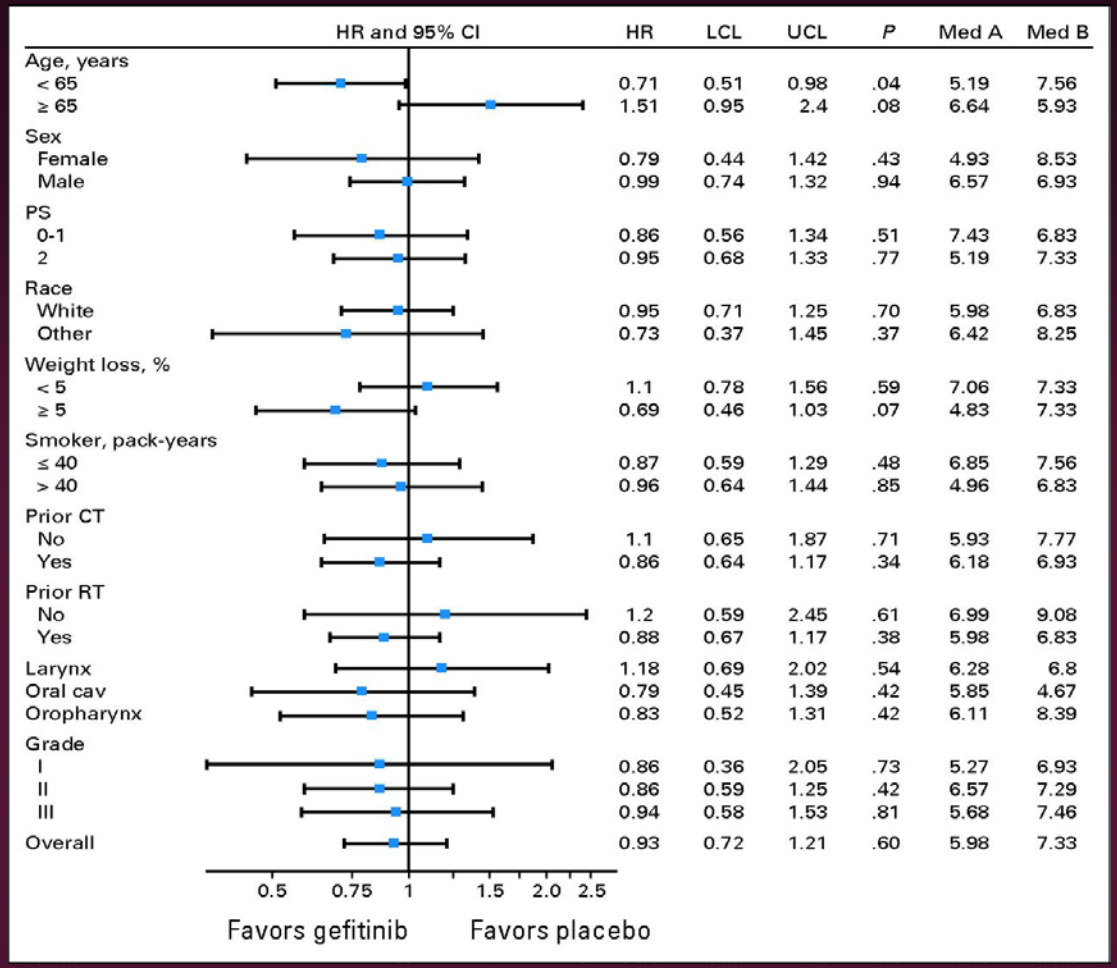
Subgroup Analyses

- Are there particular subpopulations which demonstrate effect?

- Be cautious – do not over-interpret

- Remember, if H0 is true, $\alpha=0.05$ means 1 of 20 tests significant *by chance alone*

McMaster University
HEALTH SCIENCES



Results: In an unplanned subgroup analysis, we found that patients younger than 65 years derived survival benefit from combination therapy (median OS, 7.6 months with docetaxel/gefitinib *v* 5.2 months with docetaxel/placebo; $P = .04$).

Conclusion: Our observation of a potential survival benefit with the addition of gefitinib to docetaxel in younger but not older patients may warrant further validation in clinical studies.

NOTE: 22 tests in H&N cancer – plausible? HPV (?) Mutations (?)

McMaster University HEALTH SCIENCES

Argiris et al. JCO 31(11); April 2013, doi: 10.1200/JCO.2012.45.4272

McMaster University
HEALTH SCIENCES