Single Trial w/ 10,000 Subjects

Pooled Trials w/ 10,000 Subject-Trials

Control Database (Never Surgery)

Trial 1	Trial 2	Trial 3	Trial 4
Trial 5	Trial 6	Trial 7	Trial 8
Trial 9	Trial 10	Trial 11	Trial 12

1) Sample Covariates

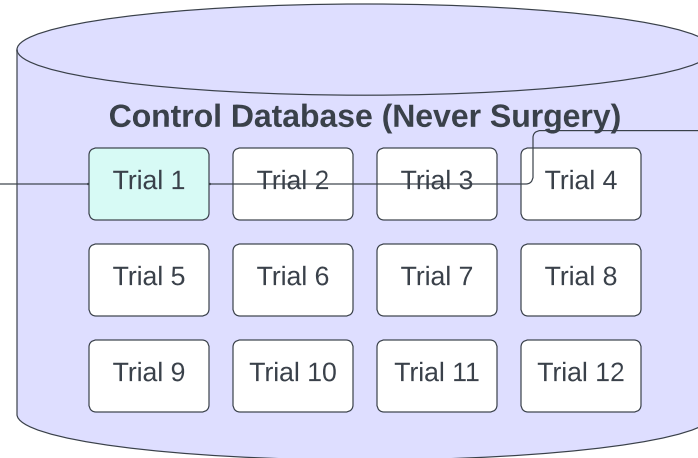Sample covariates of 10,000 subjects from Trial 1 (January 2011)

2) Sample Treatment

Randomly treat 10% of subjects w/ some binary treatment

3) Sample Outcome

Sample month by month from pooled logistic regression until failure or 60 months of follow up

1) Sample Covariates

Sample covariates of 10,000 subjects from Trial 1 (January 2011)

Note I keep the BMI covariate the same over time (may want to change this in future)

2) Sample Treatment Time

Randomly treat 10% of subjects w/ some binary treatment

For those that are treated, they get their treatment uniformily between month (trial) 1 and 12

3) Sample Outcome

Sample month by month from pooled logistic regression until failure or 60 months of follow up **the overall first trial (January 2011).** This is true even if we don't see the person.

Shift outcomes relative to first trial. If subject experiences outcome before some subsequent trial, those obvioulsy get dropped.

Thus, effective # of subject-trials is < 10,000 by a little bit

4) Sample # of Trials for Each Subject

1) Get empirical distribution of # of trials that controls are in (out of the 12)
2) Get empirical distibution of # of trials that cases are in (out of 12, w/ their case trial being the last one)

For controls: draw # of trials from 1)
For cases: draw # of trials from 2) conditional on last one being treatment time
Sample which trials they are in uniformly w/ exception that cases are in the treatment trial

5a) Truncate to 10,000 Subject trials

Randomly take subjects (and all their associated trials) until 10,000 subject trials have been added

5b) Version Where Everyone gets 12 Trials

Randomly take subjects (and all 12 associated trials) until same # of subjects (but more subject trials) as 5a)