

## Device: Sample Size for Device Non-Inferiority Trial

**Goal:** To estimate sample size for Bayesian non-inferiority assessment of a new device versus a control.

**Background:** A new orthopedic device is being developed. A non-inferiority trial is designed versus a control device where endpoint is device success rate. Sample size needs to be estimated for a Bayesian non-inferiority comparison. Historical priors are available on the control.

**Bayesian Model:** The success rate is a binary response that is analysed using an independent proportion model.

Priors were vague on the new device response rate:

$$p_1 \sim \text{beta}(0.5, 0.5),$$

and somewhat informative on the control device rate of success (6/10 data):

$$p_0 \sim \text{beta}(6.5, 4.5).$$

**Adaptive design:** This is a balanced trial. Power was evaluated for sample sizes of 200 to 290 per group. To do so, non-adaptive cohorts of size 20 (10/arm) were created in Decimaker.

**Decision:** The non-inferiority decision is based on the posterior probabilities. It is concluded when:

$$\Pr[p_1 > p_0 - 0.1] > 95\%.$$

In this case, a non-inferiority margin of 10% was selected as being clinically relevant.

**Simulation:** Power was assessed using simulations assuming that true success rate in the control and new device arms were both 60%.

Cohorts of increasing sample size: from 200 to 290/arm were analysed and the study was stopped as soon as the non-inferiority decision was met. We report the expected cumulative percentage of trial stops by cohort as power estimator.

Type 1 error rate could be estimated using the same logic assuming that  $p_1$  equals the non-inferiority margin of  $p_0 - 0.1$  (result not reported).

**Results:** One thousand studies were simulated to assess the power of the Bayesian test procedure. Results were compared to a frequentist power evaluation for the one-sided comparison of 2 proportions (50 to 60%) at the 5% significance level

Table 1. Power for the Bayesian and Frequentist Non-Inferiority Tests.

Procedure	N=200	N=210	N=220	N=230	N=240	N=250	N=260	N=270	N=280	N=290
Frequentist	0.643	0.662	0.679	0.696	0.712	0.728	0.742	0.756	0.770	0.782
Bayesian	0.665	0.71	0.743	0.778	0.796	0.815	0.828	0.845	0.85	0.869

Because it relies on prior information, the Bayesian procedure is requires fewer subjects than the frequentist non-inferiority test. To achieve a 80% power, approximately 250 patients are required /arm with the Bayesian procedure. To achieve the same power, the sample size needs to be increased to 305 patients/arm using a Frequentist approach.

**Decimaker:** The Decimaker study project may be found [here](#).

### References:

- FDA CDRH. Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials - Draft Guidance for Industry and FDA Staff. <http://www.fda.gov/cdrh/osb/guidance/1601.html>, May 2006.