A Worked Example for a *Post Hoc* Stratified Analysis of a Net Difference in a Cohort Design

The calculator will use the parameter estimates you have provided to calculate the required sample size of groups per condition for the intervention effect (Δ) and the level of power that you specified.

In addition, the calculator will use the other parameter estimates to calculate the detectable difference (Δ) that is available with the level of power that you specified as a function of the number of groups (g) in each condition and the number of members (m) in each group x stratum cell, using the formula below.

$$\Delta = \sqrt{2\times 2\times 2\left(\frac{\left(\sigma_y^2(1-ICC)\left(1-R^2_{y\cdot x_m}\right)\left(1-r_{ys_m}\right)\left(1-r_{yy_m}\right)+m\sigma_y^2(ICC)\left(1-R^2_{y\cdot x_g}\right)\left(1-r_{ys_g}\right)\left(1-r_{yy_g}\right)\right)}{mg}\right)\left(t_{\alpha/2}+t_\beta\right)^2}$$

Here σ_y^2 is the variance of the outcome variable ignoring any expected ICC, ICC is the intraclass correlation among the members in the same group, $R^2_{y\cdot x_m}$ is the proportion of variance explained by the member-level covariates, r_{ys_m} is the correlation between the outcome variable and the stratification variable at the level of the member, m is the number of members in each group x stratum cell, assumed to be one half of the value of m entered in Step 5, $R^2_{y\cdot x_g}$ is the proportion of variance explained by the group-level covariates, r_{ys_g} is the correlation between the outcome variable and the stratification variable at the level of the group, g is the number of groups or clusters in the each condition, $t_{\alpha/2}$ is the t-value selected based on the two-tailed alpha level and available degrees of freedom, and t_β is the t-value selected based on the desired power and available degrees of freedom.

As an example, consider the following set of parameter estimates:

Let m per group = 100 from Step 5. With s=2, m per group x stratum cell = 50.

Let g =46 with x_m=4 df used for member-level covariates and x_g=1 df used for group-level covariates. Then

$$df = (t-1)(s-1)c(g-1)-df_g = (2-1)(2-1)2(46-1)-1 = 89$$

For a two-tailed type 1 error rate of 5% and power of 80%

$t_{\alpha/2} = 1.9870$ and $t_\beta = 0.8457$

Let σ_y^2= 1.0 and ICC=0.05.

Let $R^2_{y\cdot x_m} = 0.20$ and $R^2_{y\cdot x_g} = 0.00$.

A Worked Example for a *Post Hoc* Stratified Analysis of a Net Difference in a Cohort Design

Let $r_{ys_m} = 0.10$ and $r_{ys_g} = 0.00$.

Let $r_{yy_m} = 0.70$ and $r_{yy_g} = 0.20$.

Then

$$\Delta = \sqrt{2 \times 2 \times 2\left(\frac{\left(1(1-0.05)(1-0.2)(1-0.1)(1-0.7)+50(1)(0.05)(1-0)(1-0)(1-0.2)\right)}{50 \times 46}\right)(1.9870+0.8457)^2}$$
$$= 0.2481$$

The detectable difference given these parameter estimates is 0.2481standard deviation units.