The Use of Power Analysis in Small Sample Pre-clinical Imaging Studies

Edward J. Soares, Ph.D.[1], Jacob Y. Hesterman, Ph.D.[2], Jack W. Hoppin, Ph.D.[2]

1. College of the Holy Cross, Worcester, MA, United States; 2. inviCRO LLC, Boston, MA, United States.

Objective

The importance of determining the correct sample size for clinical and pre-clinical studies is well established [1-3]. Power analysis (PA) is a method by which a direct relationship between effect size and sample size may be derived. The goal is to determine the sample size required to identify a certain effect size as statistically significant. Typical effect sizes include:

- Difference between group means (independent samples)
- Difference between baseline/follow-up means (matched pairs)
- % change in these differences relative to one group or time point

A z-statistic related to the standard normal distribution is often used as the basis for sample size calculations [1,4]. However, for small sample sizes typically seen in pre-clinical studies, this approach can under-estimate the number of subjects required to identify an effect size as significant [5-6]. Scheibe [7] described an equivalent approach to PA using the independent samples t-statistic. The non-centrality parameter of the t-distribution allows one to establish a direct relationship between effect size and sample size.

Here, we review and compare small- and large-sample methods for performing PA. Included are the methods for matched pairs (MP) comparisons and independent samples (IS) comparisons based on the t-distribution, as well as their large-sample approximations. Also, we demonstrate the influence of the proportionality between group variations on sample size (IS), as well as the influence of correlation between baseline and follow-up observations on sample size (MP). Finally, we quantify the % decrease in effect size when adding additional subjects (based on n=3) for both IS and MP analyses.

Methodology

Power analysis requires the identification of the following:

- Null (H_0) and alternate (H_a) hypotheses
- Test statistic and corresponding probability distribution
- A form of the non-centrality parameter (δ) for the distribution
- Significance level (α) and power (1-β)
- Either pilot data or assumptions about the other parameters of the distribution (value or relationship between standard deviations, correlation coefficient, etc.)

Computational Procedure (IS or MP power analysis):

- Choose a value for n
- Set significance level α = 0.05 and power = 0.8 (β = 0.2)
- Determine degrees of freedom for the t-distribution
- Use cumulative distribution function (CDF) of the non-central t-distribution to find the value of the non-centrality parameter δ
- Compute effect size Δμ
- Plot effect size Δμ versus sample size n

Examples of non-centrality parameters:

$$\delta = \frac{\Delta\mu}{\sigma\sqrt{2/n}} \quad \text{IS (equal vars)}$$

$$\delta = \frac{\Delta\mu}{\sqrt{\sigma_1^2 + \sigma_2^2/\sqrt{n}}} \quad \text{IS (unequal vars)}$$

$$\delta = \frac{\Delta\mu}{\sigma_d/\sqrt{n}} \quad \text{MP}$$

Summary

- As variability in the second sample increases relative to the first (IS test), the number of subjects required to identify a fixed effect size as significant increases.
- As correlation between the two measurements decreases (MP test), the number of subjects required to identify a fixed effect size as significant increases.
- For small sample sizes, use of the z-statistic will underestimate the n required to find the effect size statistically significant.
- Adding two additional subjects decreases effect size by 35% (IS test) and decreases effect size by 49% (MP test).

Contact Info: For questions, contact Dr. Soares (esoares@holycross.edu)

Results: Effect of Unequal Variability (Independent Samples) and Correlation (Matched Pairs)

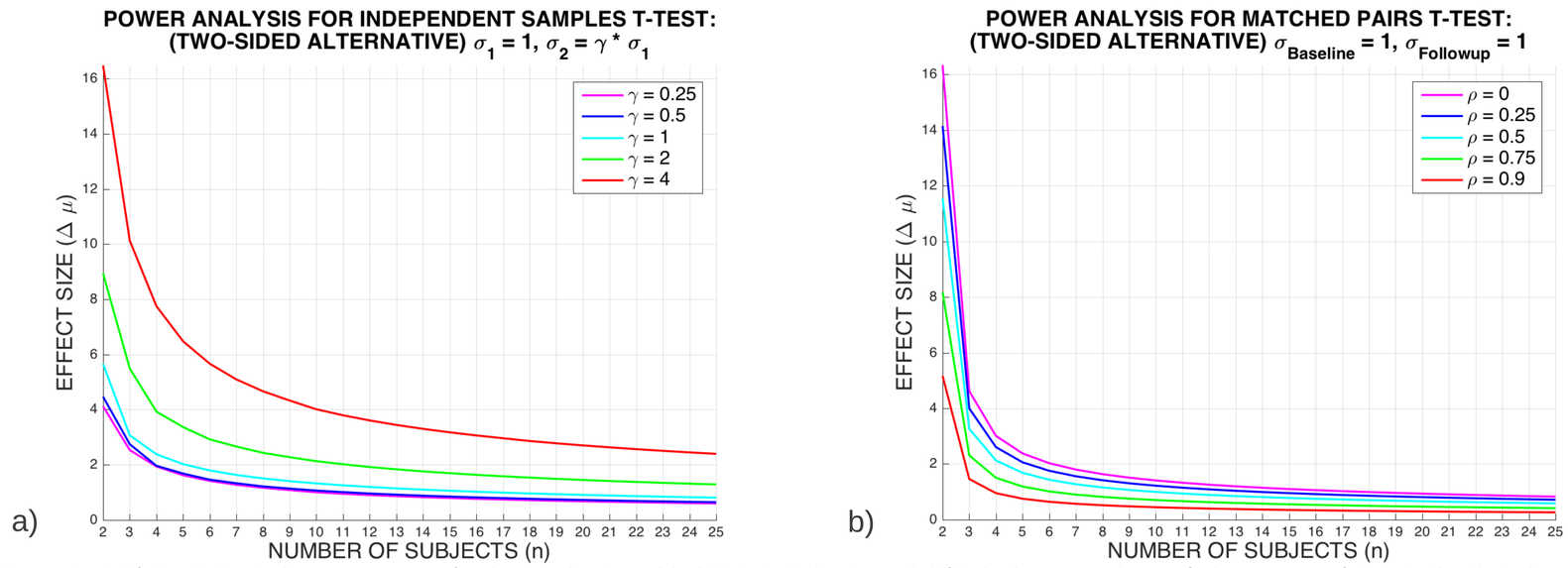

Figure 1. a) Plots of effect size versus sample size for the two-sided IS t-statistic. As variability in the second sample increases relative to the first, the number of subjects required to identify a fixed effect size as significant increases; b) Plots of effect size versus sample size for the two-sided MP t-statistic. As correlation between the two measurements decreases, the number of subjects required to identify a fixed effect size as significant increases.

Results: Comparison of Large vs. Small Sample Methods

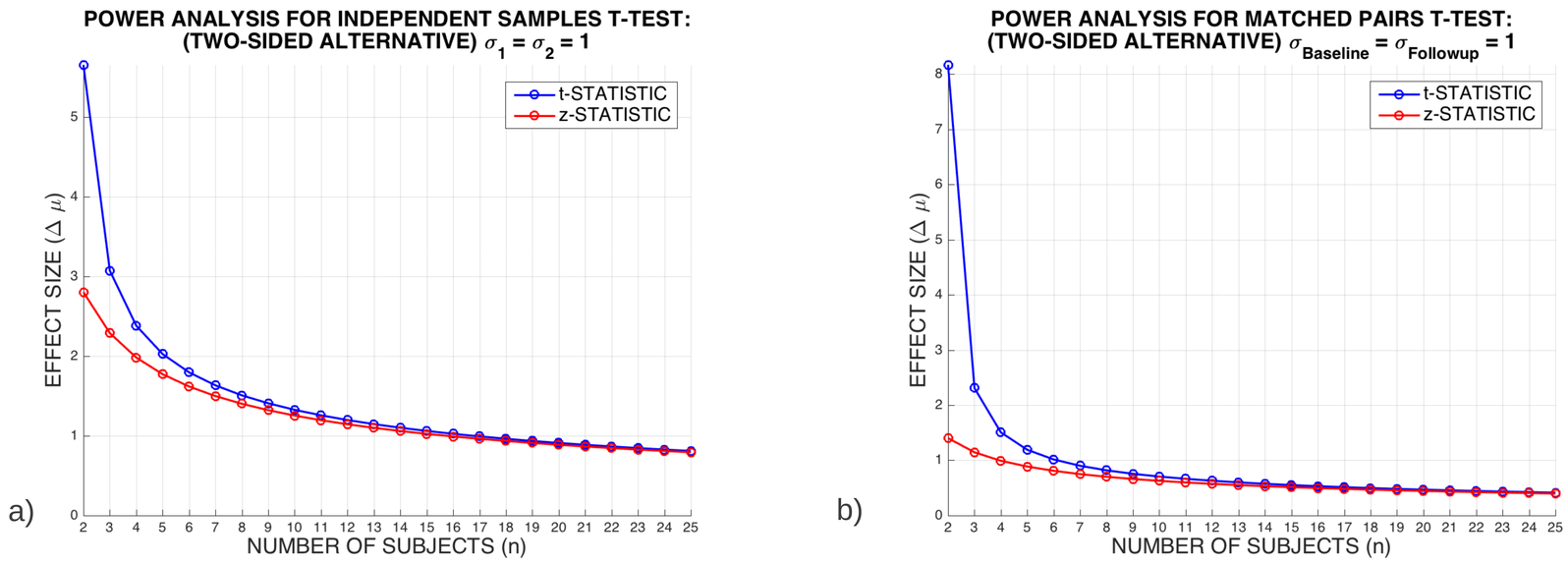

Figure 2. a) Plots of effect size versus sample size for the two-sided IS t- and z-statistics; b) Plots of effect size versus sample size for the two-sided MP t- and z- statistics. In both cases, for small sample sizes, use of the z-statistic will underestimate the n required to find the effect size statistically significant.

Results: Percent Decrease in Effect Size Achieved By Adding Additional Subjects

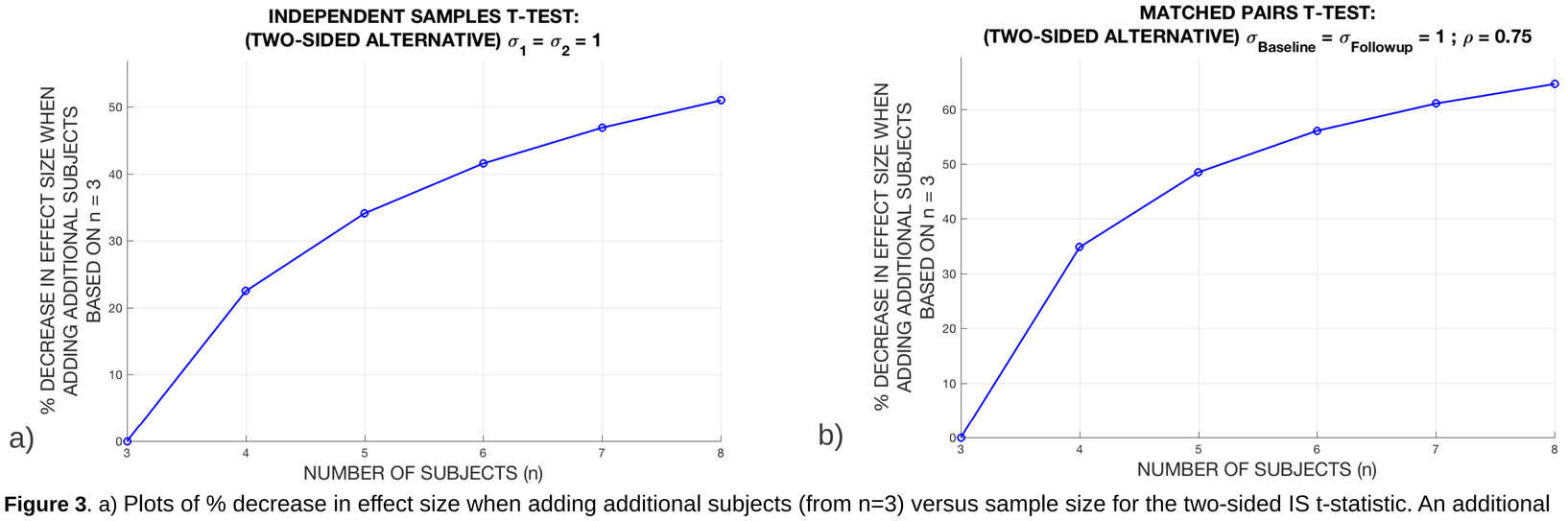

Figure 3. a) Plots of % decrease in effect size when adding additional subjects (from n=3) versus sample size for the two-sided IS t-statistic. An additional subject per group decreases effect size by 23%, while adding two additional subjects decreases effect size by 35%; b) Plots of % decrease in effect size when adding additional subjects (from n=3) versus sample size for the two-sided MP t-statistic. An additional subject per group decreases effect size by 35%, while adding two additional subjects decreases effect size by 49%.

References

[1] J Eng, Radiology, 227, 309-313, 2003; [2] Editorial, *Nucl. Med. and Biol.*, **34**, 229-232, 2007; [3] Editorial, *Nucl. Med. and Biol.*, **35**, 1-2, 2008; [4] B Rosner, *Fundamental of Biostatistics*, 7th ed., 302-305, 2011; [5] S Chow, J Shao, & H Wang, *J. Biopharm. Stat.*, **12**(4), 441-456, 2002 ; [6] DA Harrison & AR Brady, *The Stata Journal*, **4**(2), 142-153, 2004; [7] P Scheibe, *Nucl. Med. and Biol.*, **35**, 3-9, 2008.