Published in final edited form as:

Clin Trials. 2013 ; 10(6): 990–991. doi:10.1177/1740774513506965.

Letter to the Editor: An Alternative Property for Evaluating Sample Size for Normal Data using Preliminary Data

Michael P. Fay

Chen, Zhang, Lu and Chen (2013) rightly pointed out that naively treating sample variances as the true variance can lead to under-powered studies. They studied several methods for rectifying this problem and proposed a property for evaluating the methods. I would like to mention a different evaluating property and an associated sample size method previously described in Fay, Halloran, and Follmann (2007).

Ideally, we want the probability of rejection given the preliminary data and given a sample size function of that data to be equal to the planned power, say 80%. We cannot calculate that power exactly since we do not know the variance from the preliminary data; however, under the normality assumption, we can estimate the distribution of that power, and set properties for evaluating sample size functions using that distribution.

For clarity, we give notation and define the properties described in Chen, *et al.* (2013) and Fay *et al.* (2007). Let $Pow(n_{new}, \sigma^2)$ be the power given n_{new} is the fixed sample size per group for the new study, and σ is the known standard deviation. Chen, *et al.* (2013) call this the actual power, while Fay *et al.* (2007) call it the fixed-n power. We do not observe σ^2 , but only observe s^2 , an estimate of σ^2 ; therefore, we cannot calculate $Pow(n_{new}, \sigma^2)$ directly. Suppose $\eta(\sigma^2)$ is a function of the variance which returns the per-group sample size. The usual normal approximation sample size function is

$$\eta_{norm}(\sigma^2) = \left(\frac{2(Z_{\alpha/2} + Z_{\beta})^2}{\delta^2} \right) \sigma^2 \quad (1)$$

where $\alpha/2$ is the one-sided significance level, $1 - \beta$ is the planned power, δ is the clinically significant difference, and Z_q is the $(1 - q)$ th quantile from the standard normal distribution. So that

$$Pow(\eta_{norm}(\sigma^2), \sigma^2) \approx 1 - \beta,$$

but

$$Pow(\eta_{norm}(s^2), \sigma^2) \approx \Phi \left(-Z_{\alpha/2} + (Z_{\alpha/2} + Z_{\beta}) \sqrt{\frac{s^2}{\sigma^2}} \right),$$

where s^2 is estimated with d degrees of freedom, and the approximations come from using the Z -test instead of the t -test. Since the preliminary data are normally distributed, dS^2/σ^2 is

chi-square with d degrees of freedom, where S is the random variable associated with s . So we can calculate properties of $Pow(\eta_{norm}(S^2), \sigma^2)$ knowing only d .

Here are the two evaluating properties for a sample size function η with 80% planned power:

Chen, et al.: We want 80% of $Pow(\eta(S^2), \sigma^2)$ to be greater than or equal to 80%.

Fay et al.: We want to reject at least 80% of the time when we use η for sample size calculations. This is equivalent to $E_S [Pow(\eta(S^2), \sigma^2)] \geq 0.80$, where Fay et al. (2007) call the left-hand-side of that expression the *overall* power.

It seems to me that the latter property captures more naturally what we want from a sample size function.

To adjust η_{norm} to have overall power approximately equal to the planned power, you increase $1 - \beta$ in equation 1 to the calibrated power, $1 - \beta^*$, which is based on d . For example, if $d = 15$, $\alpha = 0.05$, and the planned power is $1 - \beta = 0.80$, then using $1 - \beta^* = 0.8426$ for $1 - \beta$ in equation 1 will give overall power of about 80%. If we have multiple preliminary studies, we can use this method with the weighted estimate of variance,

$$\frac{\sum_{i=1}^K (n_i - 1) S_i^2}{\sum_{i=1}^K (n_i - 1)},$$

since it is distributed chi-square with degrees of freedom equal to $\sum_{i=1}^K (n_i - 1)$. For details see Fay et al. (2007) and the R package *ssanv*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Chen H, Zhang N, Lu X, Chen S. Caution regarding the choice of standard deviations to guide sample size calculations in clinical trials. *Clinical Trials*. 2013; 10:522–529. [PubMed: 23794405]
- Fay MP, Halloran ME, Follman DA. Accounting for variability in sample size estimation with applications to nonadherence and estimation of variance and effect size. *Biometrics*. 2007; 63:465–474. [PubMed: 17688499]