

Sample Size Calculation for a SARS-Cov-2 Serological Test Validation Study

Enveritas

April 26, 2020

We need exact methods because the normal approximation is not accurate at our typical required sample size (between 150 and 250 patients) when the efficacy parameter is expected to be 0.95 or higher. This is particularly serious because the comparison of exact methods with normal approximations generally shows that the normal approximation significantly underestimates the required sample size in these situations. So not only is it inaccurate but it also errs in precisely the wrong direction!

Diagnostic tests are typically associated with very high values of sensitivity and specificity. In such cases, sample size calculation based on a normal approximation to the binomial distribution may be misleading. With modern computing, it is feasible to replace normal approximations with exact methods.(Chu & Cole, 2007)

for tests of a single binomial proportion compared to a hypothesized standard, statistical power does not monotonically increase as a function of increasing sample size: the power function appears to be “saw toothed”(Chu & Cole, 2007)

To evaluate the accuracy of specificity estimate, the experimenter must further estimate some confidence limit. The $1 - \alpha$ lower confidence limit for Sp can be thought of as the lowest value of Sp that is not rejected by a one-sided test of level α of the null hypothesis $Sp = Sp_L$ against the alternative hypothesis $Sp > Sp_L$. Upper confidence limits are defined in an analogous manner, but are irrelevant here, because the concern is that the test actually performs worse, not better, than indicated by the observed proportion of positive outcomes in the trial sample.(Flahault, Cadilhac, & Thomas, 2005). The normal approximation is explained in (Machin, Campbell, Fayers, & Pinol, 1997) and (Hajian-Tilaki, 2014).

$$N = \frac{\left(z_{1-\alpha/2}\sqrt{\pi_1(1-\pi_1)} + z_{1-\beta}\sqrt{\pi_2(1-\pi_2)}\right)^2}{\delta^2} \quad (1)$$

one sided test (tail), reordering and setting $\pi_1 = \pi - \delta$ and $\pi_2 = \pi$:

$$N = \frac{\left(z_{1-\beta}\sqrt{\pi(1-\pi)} + z_{1-\alpha}\sqrt{(\pi-\delta)(1-\pi+\delta)}\right)^2}{\delta^2} \quad (2)$$

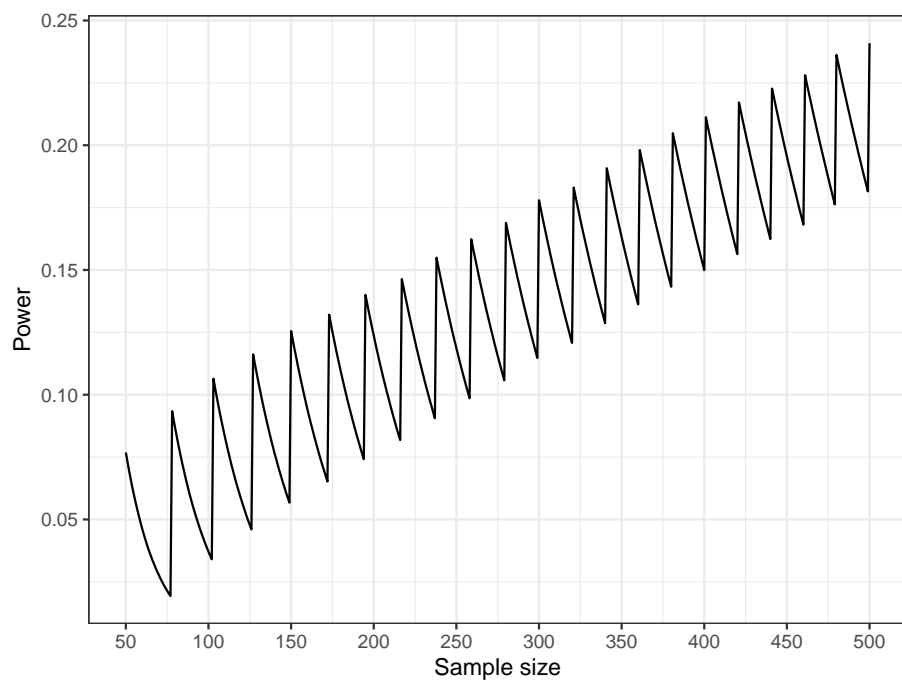


Figure 1: power as function of increasing sample size does not increase monotonically

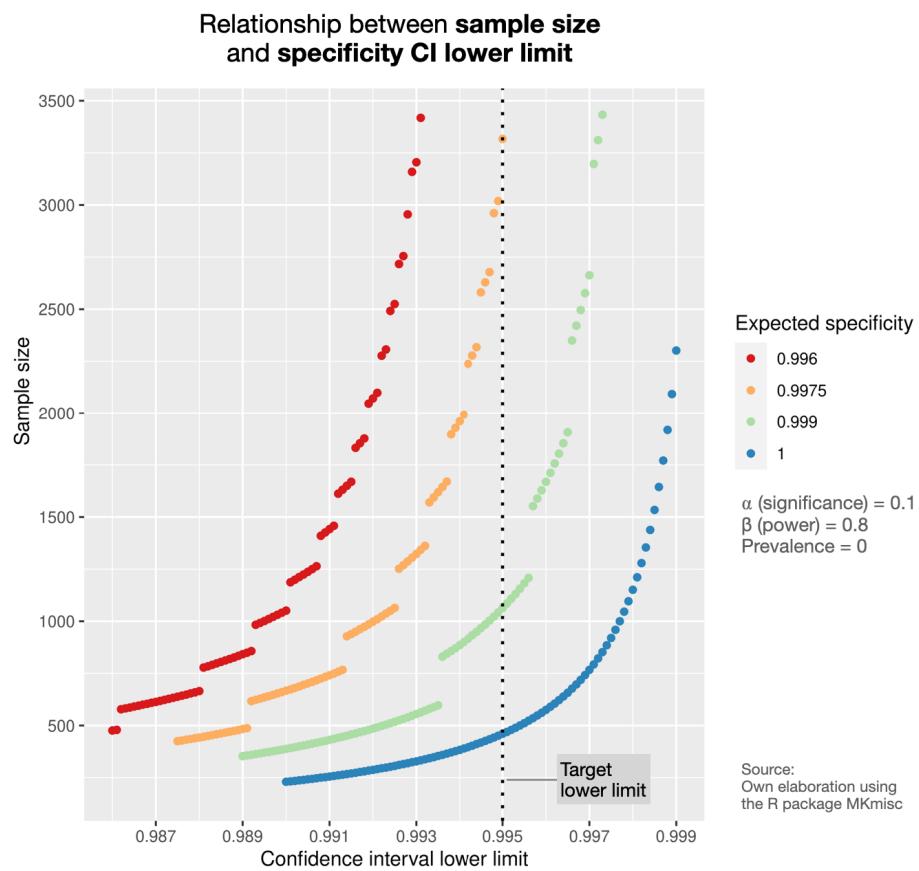


Figure 2: lim vs ss

$$\sum_{i=x}^n \binom{n}{i} \hat{\pi}_L^i (1 - \hat{\pi}_L)^{n-i} = \alpha \quad (3)$$

$$\sum_{i=q_\beta}^n \binom{n}{i} \gamma^i (1 - \gamma)^{n-i} = \alpha \quad (4)$$

References

- Arslan, A. K., Yaşar, Ş., Çolak, C., & YOLOĞLUa, S. (2018). Wsspas: An interactive web application for sample size and power analysis with r using shiny. *Türkiye Klinikleri Journal of Biostatistics*, 10(3).
- Campbell, M. J., Julious, S. A., & Altman, D. G. (1995). Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *Bmj*, 311(7013), 1145–1148.
- Chernick, M. R., & Liu, C. Y. (2002). The saw-toothed behavior of power versus sample size and software solutions: single binomial proportion using exact methods. *The American Statistician*, 56(2), 149–155.
- Chu, H., & Cole, S. R. (2007). Sample size calculation using exact methods in diagnostic test studies. *Journal of Clinical Epidemiology*, 60(11), 1201–1202.
- Flahault, A., Cadilhac, M., & Thomas, G. (2005). Sample size calculation should be performed for design accuracy in diagnostic test studies. *Journal of clinical epidemiology*, 58(8), 859–862.
- Fleiss, J. L., Tytun, A., & Ury, H. K. (1980). A simple approximation for calculating sample sizes for comparing independent proportions. *Biometrics*, 343–346.
- Hajian-Tilaki, K. (2014). Sample size estimation in diagnostic test studies of biomedical informatics. *Journal of biomedical informatics*, 48, 193–204.
- Kohl, M. (2019). MKmisc: Miscellaneous functions from M. Kohl [Computer software manual]. Retrieved from <http://www.stamats.de> (R package version 1.6)
- Machin, D., Campbell, M., Fayers, P., & Pinol, A. (1997). Sample size tables for clinical studies. 1997. *Malden, MA: Blackwell Science*, 2.
- R Core Team. (2020). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>