CLINICAL TRIALS

Clinical Trials
1–2

DOI: 10.1177/17407745221110185
journals.sagepub.com/home/ctj
SAGE

On the proposed use of a finite-population correction factor in clinical trials

Andrew V Frane

Mauguen[1] proposed that when a disease is "rare," the population is finite, in which case clinical trials should use the following finite-population correction factor to adjust the estimated standard error: $c = \sqrt{(N-n)/(N-1)}$, where c is the finite-population correction factor (to be multiplied by the conventional standard error estimate), N is the population size, and n is the sample size. Typically, this adjustment is negligible,[2] as $c \approx 1$ when n is a small fraction of N. But c decreases toward 0 as n increases toward N. Consequently, if a large proportion of a finite population is sampled, c can notably reduce the estimated standard error and increase the probability of statistical significance. Although that may sound appealing, applying c as proposed is not advisable.

For one thing, a disease being rare does not imply a finite population for inferential purposes. Mauguen[1] contrasted "a population of up to 190,000 newly diagnosed patients per year" versus "a population of less than 100 patients per year."[1] But in statistical inference, the population is not generally defined by the number of patients "per year." Rather—as Mauguen rightly noted—"The sampling population is the set of individuals for which the inference is made."[1] And in clinical trials, inferences are made regarding how well the treatment will work—not only for individuals who presently have the disease but presumably also for individuals in the future who will have the disease. Thus, the population is likely to be indeterminately large, even when studying rare illnesses. Indeed, for designed experiments, such as clinical trials, the population is typically considered as effectively infinite.[3]

Furthermore, even when assuming a finite population, c is invalid for comparing randomized groups, for example in Mauguen's[1]"comparison between two treatment arms" example. For estimating a population mean using a single sample, the finite-population correction factor is indeed c. But for comparing means using two random samples of a population,[4] the finite-population correction factor is $c_2 = \sqrt{(N/(N-1))}$. Notice c_2 is >1 (negligibly, unless N is tiny), whereas c is <1. Consequently, applying c instead of c_2 when comparing randomized groups underestimates standard error (on average) by a factor of $c/c_2 = \sqrt{1-n/N}$. As I confirmed analytically (see online supplement) and by simulation, the resulting type I error rate approaches 50% for a one-sided test as n approaches N (considering n as the total sample size). Thus, the ostensible "power gains" shown in Mauguen's table and figures are meaningless because they are only achieved by applying the wrong finite-population correction factor, which underestimates standard error. Consider the behavior of c in the limit: When $n = N$, c reduces the estimated standard error to 0. That would make sense when estimating a single mean as sampling the entire population would eliminate all uncertainty. But when estimating a mean difference using randomized groups, even sampling the entire population would not eliminate the uncertainty created by random assignment.

In summary, applying c is typically inadvisable in clinical trials, even when the disease is rare. Moreover, c is invalid for comparing randomized groups, which is the type of comparison clinical trials primarily use.

Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Department of Psychology, Occidental College, Los Angeles, CA, USA

Corresponding author:
Andrew Frane, Department of Psychology, Occidental College, 1600 Campus Road, Los Angeles, CA 90041, USA.
Email: avfrane@oxy.edu

ORCID iD

Andrew V Frane https://orcid.org/0000-0002-5057-7567

Supplemental material

Supplemental material for this article is available online.

References

1. Mauguen A. Accounting for the rarity of the disease when designing clinical trials with a focus on pediatric cancers. *Clin Trials*. Epub ahead of print 1 March 2022. DOI: 10.1177/17407745221080728.
2. Snedecor GW and Cochran WG. *Statistical methods*. 8th ed. Ames, IA: Iowa State University Press, 1989.
3. Kozak M. Finite and infinite populations in biological statistics: should we distinguish them? *J Am Sci* 2008; 4: 59–62.
4. Bondy WH and Zlot W. The standard error of the mean and the difference between means for finite populations. *Am Stat* 1976; 30: 96–97.