In multicentre trials, do 95% confidence intervals of percentages need correcting for centre effect, and if yes how?

Paul Talsma [1], Francesco Innocenti [2], Mamadou Drame [3],

[1] Statistics Department, Vivos Technology Limited (Phastar), London, UK, [2]Department of Methodology and Statistics, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands, [3] Vaccines Biostatistics Research and Development Department, GlaxoSmithKline (GSK), Rockville, US

BACKGROUND

Percentages in multicentre trials are of major importance for drawing conclusions for treatment efficacy and/ or safety. They are normally accompanied by a 95% confidence interval. 3 important questions:

1.Which method should be used to construct the 95% CI?
2.Should the 95% CI be corrected for centre effect?
3.If yes, how should this be done?

METHODS

A simulation study was conducted to evaluate the respective ways of constructing a 95% CI for a percentage, with factors:

1) Using the Wilson/ score interval, Agresti-Coull and 'exact' (Clopper-Pearson).
2) The underlying (population) percentages in the centres are the same or differ by a small, medium, or large amount, using criteria from Cohen (1978).
3) Varying number of centres: 2, 4, 8, 16 and 32.
4) Varying numbers of subjects per centre: 10, 20, 40, 80, 160.
5) While keeping total n constant, varying number of subjects across centres with average ratio of centre size of 1, 1.5, 2, 2.5, 3, respectively.
6) Different ways of estimating the Intra-Cluster Correlation (ICC): ANOVA, Fleiss-Cuzick, Pearson.
7) Different ways of calculating the Design Effect.

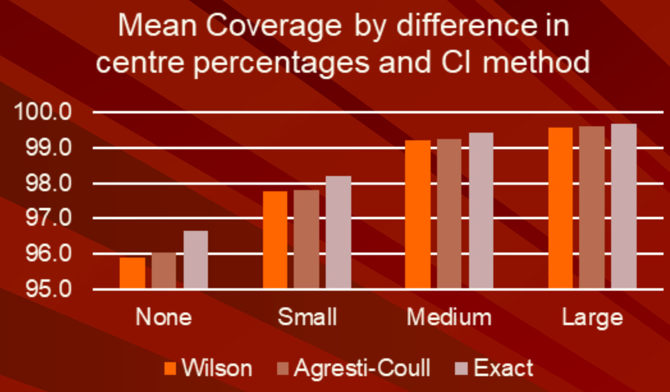
SUMMARY OF RESULTS

Most coverage percentages are above 95. Coverage is therefore generally adequate but can be too high which is undesirable.

The main driver to increasing coverage is the difference in percentages between centres.
Aggregated over all other categories the averages are 96.2, 97.9, 99.3 and 99.6% for no difference, small, medium and large differences, respectively.

In line with previous findings (Newcombe, 2013) the Wilson and Agresti-Coull intervals have very similar coverage, and both perform better than the exact confidence intervals under basically all scenarios.

The other factors did not affect mean coverage very much, but using corrections mentioned in the survey literature (factors 6 and 7) pushed coverage to >> 95%.



THE BINOMIAL SUM VARIANCE INEQUALITY (BSVI)

The variance of the sum of binomially distributed random variables (S_{bin}) will always be less than or equal to the variance of a binomial variable with the same n and p parameters:

$$Var(S_{bin}) = n * \overline{p} * (1 - \overline{p}) - ns^2$$

with $s^2 = \frac{1}{N}\sum_{i=1}^{k} n_i(p_i - \overline{p})^2$

s^2 will become larger with percentages being increasingly dissimilar, so the variance will go down, and coverage will **increase** not decrease.

Because of the BSVI, coverage can become much more than 95% which is basically undesirable for 95% confidence intervals. Users can either accept this or try to downward correct the width of the CI:

$$DE = \frac{\overline{p} * (1 - \overline{p}) - s^2}{\overline{p} * (1 - \overline{p})} = 1 - \frac{s^2}{\overline{p} * (1 - \overline{p})}$$

This Design Effect can be used to either correct the variance or calculate $n_e = n/DE$ and $r_e = r/DE$ and use these to construct the 95% CI. Since $DE \leq 1$ by definition, the confidence interval will always be equal or *smaller* than the one without correction.

The correction was implemented using the simulated data and it was shown that coverage becomes 95% on average.

CONCLUSIONS

This paper has taken a novel approach looking at coverage and interval width of percentages using combined data from different centres, where centre percentages were generated with pre-specified differences.

Coverage increases with increasing difference in percentages between centres. This could be explained using the binomial sum variance inequality (BSVI).

Practitioners can choose to leave as is or correct downwards using the suggested correction.

In multicentre trials, do 95% confidence intervals of percentages need correcting for centre effect, and if yes how?
Do use the Wilson or Agresti-Coull interval to construct the 95% CI, and either do not correct for centre effect if coverage above 95% is not deemed a problem, or use the proposed correction based on the binomial sum variance inequality if it is.

REFERENCES

Cohen, J. (1978). Statistical power analysis for the behavioral sciences. Psychology Press.

Newcombe, R. (2013). Confidence intervals for proportions and related measures of effect size. Ghapman & Hall/ CRC.