Modified Exact Sample Size for Test Validation Studies Incorporating Adjustment for Clustered Data

G.T. Fosgate

Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843 USA

ABSTRACT

Design of epidemiologic studies for the validation of diagnostic tests necessitates accurate sample size calculations to allow for estimation of diagnostic sensitivity and specificity within a specified level of precision and with the desired level of confidence. The objective of this paper was to design and evaluate a computer algorithm for the calculation of sample sizes for diagnostic accuracy studies incorporating clustered sampling units using a beta-binomial model. A beta-binomial distribution can be used to model a proportion expected to vary across populations. The cluster-adjusted sample size is calculated as the product of the unadjusted sample size by the design effect, which is estimable after definition of the beta distribution. Design effect estimations varied depending upon the 90% limit inputs, hypothesized proportion, and evaluated number of clusters. Beta distributions representing larger variation (wider 90% probability intervals) resulted in larger design effects. Smaller number of clusters for sampling also caused larger design effects. Estimated design effects also tended to increase as the proportion approached 1. It is important to incorporate cluster adjustment to sample size calculations when designing epidemiologic studies for estimation of diagnostic accuracy in the situation of correlated data. Beta-binomial models can be used to account for clustering and design effects can be estimated by inducing beta distributions that encompass among herd variation.

INTRODUCTION

Sample size calculation is an important design aspect of diagnostic test evaluation studies (Flahault et al., 2005). It is a common situation in veterinary medicine that sampling units are clustered in herds and within herds by management units such as pens and milking strings. Such clustering of animals should be incorporated in variance estimates when calculating confidence intervals. Sample size estimates should be based on the same statistical principles that will be employed during the analysis of collected data. Therefore, potential clustering of sampling units should be accounted for when estimating the appropriate study size.

Correlated sampling units in the situation of diagnostic test evaluation studies imply that the sensitivity and specificity of a particular test are not equal among clusters. Within each herd (or other cluster) the sensitivity and specificity is assumed to be the same but is expected to be different among clusters in an amount dependent upon the intraclass correlation. A convenient model for these data (correlated proportions) is the beta-binomial (Ridout et al., 1999). Using this model the cluster-level probability (e.g. sensitivity) is modeled by the continuous beta distribution and within each cluster units follow a binomial distribution with the same probability. The objective of the paper reported here was to design a computer algorithm for calculation of sample sizes for diagnostic accuracy studies incorporating clustered sampling units using a beta-binomial model.

MATERIALS AND METHODS

Cluster-level variability specification

The sample size routine assumes that data arise from a beta-binomial distribution. Probability of correct classification by the test (sensitivity or specificity) is assumed to follow a binomial distribution within each cluster. The accuracy among clusters is assumed to follow a beta distribution. The beta distribution used in the sample size calculation is induced through the specification of the assumed proportion and limits associated with a 90% probability interval. The assumption is made that the beta distribution over this interval can be approximated by a similar-length segment of a normal curve. This assumption allows the determination of parameters of the beta distribution based on the standard formula for calculating a

confidence interval (Daniel, 1999) and the formula for the standard deviation of a beta distribution (Agresti, 2002).

Design effect estimation

The design effect (DE) is the variation of the sampling protocol above the expected variation if simple random sampling had been employed (Bohning and Greiner, 1998). Formulas for calculating the design effect are available in the literature (McDermott et al., 1994). The beta-binomial assumption also allows the definition of among cluster variability by using a formula for the intraclass correlation associated with a beta distribution (Ridout et al., 1999).

The average cluster size (m; samples collected per cluster) and therefore the total sample size can then be solved by providing the number of clusters (k), effective sample size (ESS), and using the intraclass correlation (•) calculated from the induced beta distribution. The estimation of the cluster sample size provides all the necessary information for calculation of the design effect and the formula for calculating the average cluster size is included below.

$$m = \frac{ESS - \mathrm{r}(ESS)}{k - \mathrm{r}(ESS)}$$

Algorithm development

The cluster sample size algorithm was written in FORTRAN based on a previously published modified exact sample size routine (Fosgate, 2005). The inputs of the algorithm are the hypothesized sensitivity or specificity (expected mean over all clusters), desired error limit of estimate, desired level of confidence, number of clusters to be sampled, and the upper and lower 90% limits for the beta distribution modeling the sensitivity or specificity among clusters. The modified exact sample size is calculated and the usual normal approximation method using a standard formula (Greiner and Gardner, 2000) is also presented for comparison. Parameters of the beta distribution are induced leading to estimation of the design effect and the modified exact sample size is used as the effective sample size, or the number of sampling units that would be necessary under the assumption of independence (assumption of no clustering). The cluster adjusted sample size is then calculated as the design effect multiplied by the modified exact sample size.

RESULTS

Beta distributions estimated by the algorithm had 90% probability distributions similar to specified limits over evaluated ranges of proportions (Table 1). However, proportions closer to the boundary value of 1 resulted in induced beta distributions to be asymmetrical around the mean and wider than input specifications. Design effect estimations varied depending upon the expected sensitivity or specificity, 90% limit inputs, and number of clusters. Larger design effects resulted from beta distributions representing larger variation (wider 90% probability intervals). Increasing the number of clusters for sampling tended to decrease estimated design effects. Design effects were larger for proportions closer to 1. The effective sample size approaches an upper limit for fixed values of intraclass correlation and number of clusters (Figure 1).

CONCLUSION

The algorithm was developed specifically to aid in the design of diagnostic test evaluation studies but will function equally well for estimation of sample sizes associated with any population proportion. The implicit assumption when applying cluster adjusted sample size techniques for diagnostic tests is that the sensitivity and specificity are truly different (beyond sampling variability) among clusters and that the objective of the study is to estimate one overall population average for the accuracy measures. If the assumption is made that the sensitivity and specificity do not vary by cluster other than due to sampling variation then cluster adjusted techniques are not necessary.

It is important to incorporate cluster adjustment to sample size calculations when designing epidemiologic studies for estimation of diagnostic accuracy and other population proportions in the situation of correlated data. Beta-binomial models can be used to account for clustering and design effects can be estimated by inducing beta distributions that encompass expected among herd variation.

Table 1 True characteristics of beta distributions induced by the algorithm to approximate variation in proportions among herds and estimate design effects.

Proportion	90% range input	Estimated beta parameters	Distribution characteristics*	
			Mean	90% limit
0.50	0.48, 0.52	845.03, 845.03	0.50	0.48, 0.52
	0.45, 0.55	134.78, 134.78	0.50	0.45, 0.55
	0.40, 0.60	33.32, 33.32	0.50	0.40, 0.60
0.60	0.58, 0.62	973.45, 648.96	0.60	0.58, 0.62
	0.55, 0.65	155.25, 103.50	0.60	0.55, 0.65
	0.50, 0.70	38.36, 25.57	0.60	0.50, 0.70
0.70	0.68, 0.72	993.65, 425.85	0.70	0.68, 0.72
	0.65, 0.75	158.39, 67.88	0.70	0.65, 0.75
	0.60, 0.80	39.07, 16.75	0.70	0.60, 0.80
0.80	0.78, 0.82	865.02, 216.25	0.80	0.78, 0.82
	0.75, 0.85	137.73, 34.43	0.80	0.75, 0.85
	0.70, 0.90	33.83, 8.46	0.80	0.69, 0.89
0.90	0.88, 0.92	547.00, 60.78	0.90	0.88, 0.92
	0.85, 0.95	86.76, 9.64	0.90	0.85, 0.94
	0.80, 0.99	21.02, 2.34	0.90	0.78, 0.98
0.95	0.93, 0.97	304.29, 16.02	0.95	0.93, 0.97
	0.90, 0.99	47.89, 2.52	0.95	0.89, 0.99
	0.85, 0.99	11.26, 0.59	0.95	0.82, 1.00

*True mean and 90% probability interval of induced beta distribution.

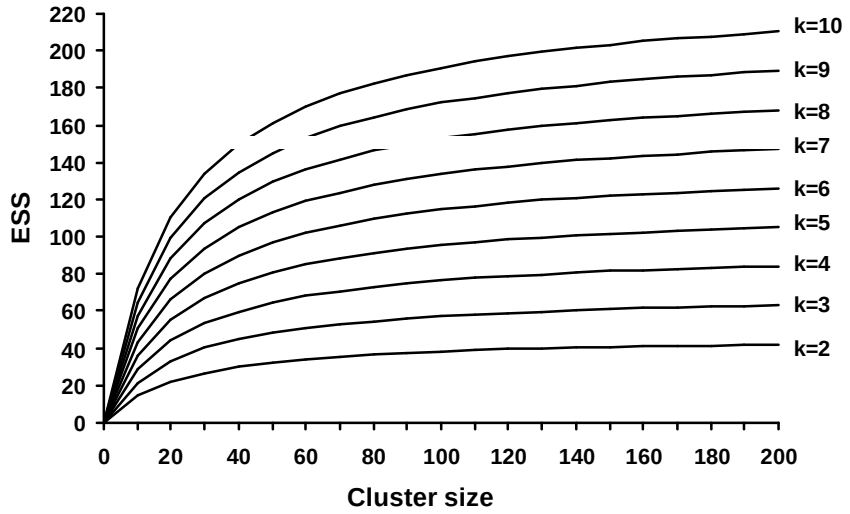

Figure 1 The consequence of increasing cluster size on effective sample size (ESS) for several number of clusters (k). Calculations based on an intraclass correlation of 0.0428 and the effective sample size of 148, which corresponds to estimating a sensitivity or specificity of 0.90 with an error of 0.05 and 95% confidence, is denoted by the dashed line.

REFERENCES

Agresti, A., 2002, Categorical data analysis, 2nd Edition. Wiley-Interscience, New York, p. 606.

Bohning, D., Greiner, M., 1998, Prevalence estimation under heterogeneity in the example of bovine trypanosomosis in Uganda. Prev. Vet. Med. 36, 11-23.

Daniel, W.W., 1999, Biostatistics : a foundation for analysis in the health sciences, 7th Edition. Wiley, New York, p. 156.

Flahault, A., Cadilhac, M., Thomas, G., 2005, Sample size calculation should be performed for design accuracy in diagnostic test studies. J. Clin. Epidemiol. 58, 859-862.

Fosgate, G.T., 2005, Modified exact sample size for a binomial proportion with special emphasis on diagnostic test parameter estimation. Stat. Med. (in press).

Greiner, M., Gardner, I.A., 2000, Epidemiologic issues in the validation of veterinary diagnostic tests. Prev. Vet. Med. 45, 3-22.

Ridout, M.S., Demetrio, C.G., Firth, D., 1999, Estimating intraclass correlation for binary data. Biometrics 55, 137-148.