Preventive Veterinary Medicine 96 (2010) 272–275

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

Short communication

Intraherd correlation coefficients and design effects for bovine viral diarrhoea, infectious bovine rhinotracheitis, leptospirosis and neosporosis in cow–calf system herds in North-eastern Mexico

J.C. Segura-Correa[a,*], D. Domínguez-Díaz[b], R. Avalos-Ramírez[b], J. Argaez-Sosa[c]

[a] Facultad de Medicina Veterinaria y Zootecnia, Campus de Ciencias Biológicas y Agropecuarias de la Universidad Autónoma de Yucatán, Km 15.5 carretera Mérida-Xmatkuil, Mérida, Yucatán, Mexico
[b] Facultad de Medicina Veterinaria y Zootecnia, Campus de Ciencias Agropecuarias de la Universidad Autónoma de Nuevo León, Escobedo, Nuevo León, Mexico
[c] Facultad de Matemáticas, Campus de Ingenierías y Ciencias Exactas de la Universidad Autónoma de Yucatán, Anillo Periférico Norte, Tablaje Cat. 13615, Colonia Chuburná Hidalgo Inn, Mérida, Yucatán, Mexico

ARTICLE INFO

Article history:
Received 30 November 2009
Received in revised form 11 July 2010
Accepted 12 July 2010

Keywords:
Intraclass correlation
Sample size
Infectious diseases
Beef cattle
Mexico

ABSTRACT

Knowledge of the intraherd correlation coefficient (ICC) and design (*D*) effect for infectious diseases could be of interest in sample size calculation and to provide the correct standard errors of prevalence estimates in cluster or two-stage samplings surveys. Information on 813 animals from 48 non-vaccinated cow–calf herds from North-eastern Mexico was used. The ICC for the bovine viral diarrhoea (BVD), infectious bovine rhinotracheitis (IBR), leptospirosis and neosporosis diseases were calculated using a Bayesian approach adjusting for the sensitivity and specificity of the diagnostic tests. The ICC and *D* values for BVD, IBR, leptospirosis and neosporosis were 0.31 and 5.91, 0.18 and 3.88, 0.22 and 4.53, and 0.11 and 2.68, respectively. The ICC and *D* values were different from 0 and *D* greater than 1, therefore large sample sizes are required to obtain the same precision in prevalence estimates than for a random simple sampling design. The report of ICC and *D* values is of great help in planning and designing two-stage sampling studies.

1. Introduction

Livestock populations are commonly clustered into herds or flocks. Estimation of prevalence in cattle or other species depends on the sampling design used, basically simple random sampling or two-stage random sampling design. Simple random sampling design commonly requires a list of each animal to be studied, which is not practically possible in the study of animal populations in large regions. However, it is possible to obtain the list of farms in the region, from which some farms could be selected and within these a sample of animals could be measured. Animals within farms do not normally keep the condition of independence between observations, which assumes that the presence or absence of disease in an animal is independent of the presence or absence of disease in another animal of the same herd; because in a same herd, animals are subject to similar management and microclimate, and many times they have similar genetic basis. This makes animal response to certain disease correlated and the response of each animal not independent of the herd. As a result of correlation, the variance in a cluster sampling may be smaller than that for a simple random sampling design of the same sample size, and in consequence the effective sample size is smaller than it should be. The loss of accuracy for using a cluster or two-stage sampling design instead of a simple random sampling is known as the design effect, *D* (Bennett et

* Corresponding author. Tel.: +52 999 423 200; fax: +52 9999 423 205.
E-mail address: segura52@hotmail.com (J.C. Segura-Correa).

0167-5877/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.prevetmed.2010.07.006

Table 1
Basic information of parameters used by the Bayesian models to estimate the intraherd correlation and design effects for four infectious diseases.

Infection	Individual prevalence (%)	Herd prevalence (%)	Specificity and sensitivity of the tests (%)	Beta parameters	Gamma parameters
Bovine viral diarrhoea	30.0	70.8	97.9, 99.7	5.21, 11.71	0.87, 0.013
Infectious rhinotracheitis	42.8	85.4	97.4, 92.4	5.92,7.81	0.606, 0.004
Leptospirosis	28.4	81.2	100.0, 99.4	7.64,19.13	1.740, 0.175
Neosporosis	11.6	81.2	100.0, 93.3	7.47,52.68	1.150, 0.020

al., 1991) or inflation factor (McDermott and Schukken, 1994).

Also, no accounting for dependence or correlation of animal results within herds when calculating sample size for a cluster or two-stage sampling design causes the overestimation of the precision of the prevalence estimates. To adjust for the loss in precision of a two-stage sampling design, the sample size must be greater. Some authors (McDermott and Schukken, 1994; Otte and Gumm, 1997; Segura-Correa and Solorio-Rivera, 2006) have reported intraherd correlations coefficients (ICC) and design effects (D) for some diseases in cattle. However, most of the studies using cluster designs do not report the ICC or D values. The lack of ICC or D estimates limits the design of cluster studies.

The ICC for the same disease varies from study to study, depending on the number of clusters and number of animals sampled within each cluster, the prevalence of the disease and the method used for their calculation (McDermott and Schukken, 1994; Segura-Correa and Solorio-Rivera, 2006), hence the need to obtain estimates for each region. Also, ICC is biased downwards if sensitivity and specificity of the diagnostic test are not 100% (Branscum et al., 2005). Couple with the effect that ICC and D have on sample size and precision of prevalence estimates, the magnitude of ICC provides biological information on the infectious agent. There are, few papers published on ICC in the Mexican context, especially for health animals surveys where herd is the primary sampling unit. However, no reports exist under the conditions of North-eastern Mexico. The aim of this study was to estimate ICC and D for BVD, IBR, leptospirosis and neosporosis diseases adjusting for the sensitivity and specificity of the diagnostic tests in cow–calf system herds in North-eastern, Mexico.

2. Materials and methods

2.1. Source of data and study design

A two-stage cross-sectional study was carried out from August 2006 to July 2007 in Nuevo Leon, Mexico. Information on 813 animals randomly selected from 48 farms was used. Herd sizes varied from 25 to 435 animals. The number of animals sampled within each herd varied between 5 and 33, and only animals over 6 months of age were sampled. The detection of serum antibodies for the viral diseases IBR and BVD, and the protozoa *Neospora caninum* was carried out using ELISA commercial kits. The detection of antibodies against different types of leptospira serovars was carried out by micro-agglutination test. The individual and herd prevalences for BVD, IBR, leptospirosis and neosporosis are shown in Table 1. The sensitivity (Se) and specificity (Sp) of the diagnosis tests for those infections were 97.9 and 99.7% (Solís-Calderón et al., 2005), 97.4 and 92.4% (Woodbine et al., 2009), 100 and 99.4% (Cho et al., 1989) and 100 and 93.3 (Wapenaar et al., 2007), respectively.

2.2. Intraherd correlation coefficient and design effect estimates

The binomial data (Y_{ij}) were modelled as beta-binomial and independent beta prior distributions were assumed for Se, Sp, μ (mean prevalence distribution of a given infection) and γ (variability of prevalence) modelled using a gamma prior distribution as described by Branscum et al. (2005):

$$Y_{ij}|p_i \sim \text{Bernoulli}(p_i)$$

$$p_i = \pi_i \text{Se} + (1-\pi_i)(1-\text{Sp})$$

$$\pi_i = \pi_i^* \quad \text{with} \quad \text{probability} \quad 1-\tau$$

$$\pi_i = 0 \quad \text{with probability} \quad \tau$$

$$\pi_i^*|\mu, \gamma \sim \text{Beta}(\mu\gamma, \gamma(1-\mu))$$

$$\mu \sim \text{Beta}(\alpha_\mu, \beta_\mu), \qquad \gamma \sim \text{Gamma}(\alpha_\gamma, \beta_\gamma)$$

$$\text{Se} \sim \text{Beta}(\alpha_{\text{Se}}, \beta_{\text{Se}}), \qquad \text{Sp} \sim \text{Beta}(\alpha_{\text{Sp}}, \beta_{\text{Sp}})$$

where p_i is the true prevalence for ith herd, π_i is the infection prevalence of the ith herd and τ is the proportion of infected herds sampled, which was set to the herd prevalence of the sample.

The ICC based on above model was calculated as suggested by Branscum et al. (2005):

$$\text{ICC} = \frac{1-\mu+\mu(\gamma+1)(1-\tau)}{1-\tau\mu}\left(\frac{1}{\gamma+1}\right)$$

D was estimated as (Bennett et al., 1991):

$$D = 1 + (b-1)\text{ICC}$$

where b is the average number of animals sampled per herd.

The model was fit using the WinBUGS program (Lunn et al., 2000). We also calculated the ICC and D and 95% credible intervals. In this study a total of 20,000 samples of possible ICC and D values were generated and the results of the first 500 rounds were deleted. The parameters of the prior beta distributions (α_μ and β_μ)