Epidémiol. santé anim., 1997, 31-32

EFFECT OF DIFFERENT SAMPLING TECHNIQUES ON ODDS RATIO ESTIMATES USING HOSPITAL-BASED CASES AND CONTROLS

Doherr M.G.[1], Carpenter T.E.[1], Wilson W.D.[1], Gardner I.A.[1]

Les biais liés à l'utilisation des données d'admission dans les hôpitaux sont rarement discutés dans la littérature vétérinaire. Les enregistrements effectués sur les patients de l'Hôpital de l'enseignement de la Médecine vétérinaire permettent d'analyser les effets de l'échantillonnage sur les estimations des odds ratio. Des chevaux atteints d'abcès à Corynebacterium pseudotuberculosis (134) sont extraits de la base de données cliniques. La population d'étude est identifiée. Plusieurs échantillons, trois randomisés, un apparié et trois avec des diagnostics différents sont sélectionnés à partir de la population d'étude. Les ratios définis comme le rapport de la proportion d'un facteur de risque calculé sur l'un des différents échantillons sur la proportion du même facteur dans la population de l'étude, sont calculés pour les quatre facteurs de risque étudiés (âge, race, sexe et type d'admission). Les ratios obtenus avec les échantillons appariés et de deux différents diagnostics ont des valeurs très étendues entraînant un biais considérable dans l'estimation des OR par rapport aux échantillons aléatoires simples et systématiques. Pour les trois techniques d'échantillonnage aléatoires, on a répété la procédure de sélection pour décrire la distribution des ratios. L'analyse de la variance et de la covariance montre que l'échantillonnage aléatoire simple et systématique donne des distributions de moyenne proche de 1 et de faible variance. Ces deux types d'échantillonnage entraînent des biais faibles dans l'estimation des OR et peuvent donc être recommandés pour la sélection du groupe témoin.

INTRODUCTION

Potential biases introduced by the application of different sampling techniques on hospital admission records have rarely been discussed in the veterinary literature. Slater at al. (1991) compared five strategies for selecting hospital controls in a canine osteochondritis dissecans example, and observed the most variation and least precision in the odds ratio estimates for risk factors with a larger number of categories.

The Veterinary Medical Teaching Hospital (VMTH) patient record database of the University of California, Davis (UCD) provided an unique opportunity to perform an in-depth analysis on the manner in which different control selection techniques do influence the risk factor estimates in a retrospective case-control study. Corynebacterium pseudotuberculosis infection, one of the commonly diagnosed infectious diseases of horses in California (Aleman et al., 1996), was used as the model disease. The main objective of this study was to describe the effect of seven different control selection techniques applied to the study base (population) on the odds ratio (OR) estimates of C. pseudotuberculosis infection in horses seen by VMTH veterinarians between July 1, 1992 and June 30, 1994. Admission type, age, sex, and breed of the horse patient, readily available from the computerized medical records, were used as indicator variables.

MATERIALS AND METHODS

Clinically confirmed cases of the disease occurring between July 1, 1992 and June 30, 1994 (n = 134) were identified from the medical record database. Inclusion criteria were the presence of a clinically confirmed abscess or a positive microbiological or serological result (antibody titer ≥1:80). All non-case horses admitted during this time period served as the study population for control selection. Stratum-specific odds ratios (OR) and 95% confidence intervals (CI) were calculated by cross-tabulating the observed case frequencies with the study base frequencies for each level of the four indicator variables admission type, age, breed, and sex. These OR's were used as the 'gold standard' in later sections of the study when selected subsets of this study population were used.

Three randomized sampling techniques (simple random, stratified random and systematic sampling), matching (6-to-1 matching on date of admission), and selection of controls based on three other diagnoses ('colic', 'cuts and lacerations', and 'fractures') were selected to derive samples from the study base. Sampling ratios (SR), defined as the ratio between the proportion of control horses in the sample p_s and the proportion of horses seen in the study population p_p were calculated for each category of the variables identified above and for each of the seven sampling techniques. Sampling was repeated 10 (systematic sampling) or 1000 (simple random and stratified random sampling) times with samples sizes of 100, 200, 400 and 800 controls, and SR's were calculated for each iteration. The SR distribution within each set (created by 10 or 1000 iterations and characterized by sampling method and sample size) was determined for all risk factor categories and expressed in terms of the mean (SR Mean), standard deviation (SR SD) and absolute deviation of the SR mean from the expected value of 1.0 (SR ΔMean).

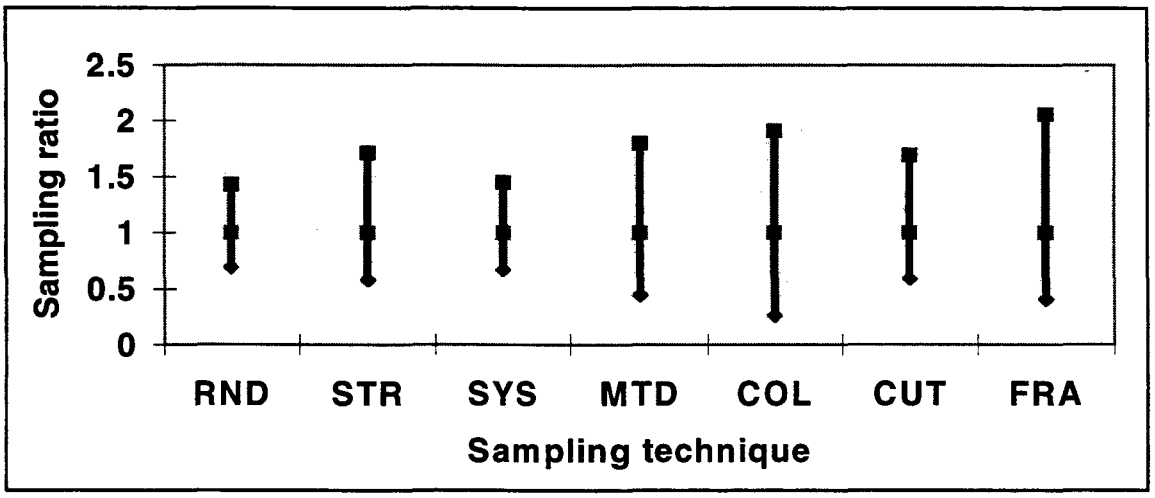
[1] School of Veterinary Medicine, Dept. of Medicine and Epidemiology, University of California, Davis, CA 95616-8737, USA

RESULTS

Observed SR ranges for all 'single' samples and all levels of the four factors under study are summarized in Figure 1. The simple random sample had the smallest range (0.74) of observed SR's (across all categories of the four variables) with a minimum of 0.69 and a maximum of 1.43. The 'fractures' sample had the largest range (0.40 - 2.06) of the seven samples, closely followed by the 'colic' sample (0.26 - 1.91).

The sampling ratios for each risk factor category were normally distributed with an expected value of 1.0. Deviations from this expected value were expressed in the distribution mean, standard distribution and absolute difference from the expected value. SR SD was significantly ($p < 0.05$) associated with the risk factor (RISKFACT), sample size (SSIZE) and number of categories within the risk factor (NOCATF). SR ΔMean was significantly associated with sampling technique (SDESIGN) and RISKFACT. In an analysis of covariance individually controlling for the effect of risk factor, SSIZE, number of iterations (NOITER) or NOCATF, SR ΔMean was always significantly ($p < 0.001$) associated with SDESIGN. Simple random sampling, regardless of the covariate used, had the smallest SR ΔMean, followed by systematic sampling and stratified random sampling.

Figure 1
Range of observed ratios between risk factor proportion in the sample and risk factor proportion in the study base (sampling ratio; expected value = 1.0) for simple random (RND), stratified random (STR), systematic (SYS), matched (MTD) sampling and the three disease 'samples' (Colic, Cuts and Lacerations, Fractures)



DISCUSSION

In our study, selection (of the study base) and misclassification bias were assumed to be identical for all control sets (since they occurred previous to the sampling procedure). This allowed us, by looking at the SR's, to estimate the effect of additional bias due to sampling design. For all sampling techniques with a random component, we in addition were able to examine the SR by drawing multiple samples from the study population. This iterative procedure had the advantage of providing information on the distribution (type and parameters) of the SR for each factor (and their category) under study.

Especially extreme sampling proportions (large deviation from population proportion) within categories of the factors admission type, age and breed were responsible for the wide range of SR's seen in the matched, colic and fracture samples. For the colic and fracture samples these extreme observations can be explained by considerable differences in distributions of these risk factors between the sample and the study base. Some of the large variation is most likely due to the small number of horses that we observed within some or all of the 14 age categories.

The fact that the variables SSIZE, RISKFACT, and NOCATF were associated with SR SD is of little surprise. An increase in sample size (with n approaching the population N) leads to more precise SR_i's (closer to the expected value of 1.0) within each iteration (number of repeated samples of size n), and a smaller SD of the distribution of derived SR_i's. RISKFACT and NOCATF effect the number of categories, i.e. the number of observations per cell used to calculate p_s and SR. The more categories (less no. of observations per category) we have, the more deviation can be expected between p_s and p_p (and therefore between SR and 1.0), which again results in a larger SD for the SR_i distribution. This is in agreement with other findings were five different control sets from hospital admissions were selected at various time points within the study period. For factors with a smaller number of categories (gender = 2, age = 6), the authors reported very little differences between the OR estimates, their confidence intervals (CI), and the significance of effects (osteochondrosis dissecans in dogs) derived from the five different control sets. Breed however, having 12 categories (and therefore a larger variation in the number of animals per category), showed OR's with considerably wider confidence intervals (less precision) (Slater et al., 1991).

Since the influence of sampling design on the dependent variables SR Mean, SD, and ΔMean was our main interest, we used the analysis of covariance with sample design as factoring and either SR Mean, SD or SR ΔMean as dependent variables to evaluate the influence of RISKFACT, SSIZE, NOITER, or NOCATF as covariates. All covariates had a significant linear relationship with SR SD ($p < 0.05$). For SSIZE, RISKFACT and NOCATF this again indicated that an increase in observations per cell or sample size significantly decreased the variability (expressed as SD) within the SR (normal) distribution. When individually adjusted for the effect of each of these covariates, SDESIGN was strongly associated with SR ΔMean. Based on these results we concluded that sampling design is influential on the absolute deviation of the SR from 1.0 (expressed as SR ΔMean), and on the SR SD. Randomized sampling procedures generally provided SR closer to 1.0 and with smaller SD than matched or different diagnosis samples, therefore introducing less sampling-related bias and variation into the OR's than other strategies. Among the randomized procedures, simple random sampling performed better than systematic or stratified sampling.

ACKNOWLEDGMENTS

This project was supported by the Center for Equine Health (CEH) with funds provided by the Oak Tree Racing Association, the State of California pari-mutuel fund, and contributions by private donors, and though a fellowship provided by the German Academic Exchange Service (DAAD) with funds from the German Federal Ministry of Education, Science, Research and Technology (BMBF).

BIBLIOGRAPHY

Aleman, M., Spier S.J., Wilson W.D. Doherr M.G., 1996. Corynebacterium pseudotuberculosis infection in horses: 538 cases (1982-1993). Journal of the American Veterinary Medical Association 209(4), 804-809.

Slater M.R., Scarlett J.M., Erb H.N., 1991. A comparison of five strategies for selecting control groups: a canine osteochondritis dissecans example. Preventive Veterinary Medicine 10, 319-326.