

## FUNDAMENTAL SAMPLING AND STATISTICAL CONSIDERATIONS FOR EPIDEMIOLOGICAL STUDIES

Lakhani K.H.<sup>1</sup>, Wood J.L.N.

*Advances in statistical and computing methods have enabled many epidemiologists to use more realistic but complex models e.g. linear and non-linear mixed models with both fixed and random effect terms. In sad contrast, the literature also contains many examples where even the basic criteria of adequate replication and statistical independence (usually due to randomisation) are not met. Inadequate replication leads to reduced statistical power and hence "not significant" test results, but the referees "allow" the authors to draw misleading inferences. Lack of statistical independence (e.g. no randomisation, clustered data, etc.) invalidates the test procedures. Erroneous inferences lead to a state of confusion in the literature, adversely influencing other researchers. Valid and trustworthy inferences are essential for maintaining excellence in epidemiology.*

### INTRODUCTION

Much of epidemiological research is based on quantitative statistical methods. The essential requirements for satisfactory and sound scientific inference are that the observed sample should represent the relevant population, that there is adequate replication and randomization, and that the analysis and interpretation of results are rigorous and reliable. Sadly, even well-known journals contain reports based on poor statistical arguments and unreliable conclusions. Fundamental statistical requirements are emphasised here.

### POPULATION

The target *statistical population* must be defined accurately. This includes its structure with respect to gender, age, weight etc. and other attributes e.g. clustering due to litters, herds, farms, etc. The *objectives* of the study should identify the relevant population factors and attributes, but also the *variables* which need to be included.

### STUDY SAMPLE

The *sample* must be *representative* of the defined population including its structure. There must be an adequate amount of *randomisation* consistent with the objectives of the study and planned significance tests. The actual level of randomisation carried out also defines the appropriate analytical model later. Adequate *replication* is essential. Studies based on small samples are liable to yield "not significant" results because of reduced *statistical power*. Such negative findings cannot provide reliable inferences about the target population.

### ANALYSIS AND INTERPRETATION

The mathematical model underlying the analysis must be appropriate and valid. The assumptions underlying the model should be testable and be reasonable both statistically and biologically. The results of the analysis should be robust to some failure of the assumptions. For example, the assumption of normality of the distribution of error is often acceptable (if replication is adequate) because of the Central Limit Theorem, but the assumption of statistical independence of the observations might be breached unacceptably if the data are repeat measurements or if the data arise from clusters of litters, herd, etc.

### EXAMPLE CASE STUDY : EFFECT OF GENDER ON DEAFNESS IN DALMATIANS

There is confusion in the published literature. One study reported an increased risk of deafness in males from a small sample "not significant" result based on only 23 dogs. Other studies have claimed no gender effect from "not significant" results. Further, even though a large sample study (900 dogs) found significantly higher prevalence of deafness in *females*, this result was arbitrarily discarded by a subsequent author who claimed that the sample size was small (!), ignoring the reported significant result. In this confused situation, a very recent study found higher incidence of deafness in females compared with that in males. These authors found that adjusting the results for the contribution of eye colour (which is thought to have an effect on deafness) exacerbated this difference. However, instead of accepting their own clear result based on 825 dogs, they concluded "The significant contribution of sex, however, was unexpected and difficult to explain". Most studies have also ignored the clustering of the observed dogs into litters. The main reason for the prevailing confusion might be that unwary referees and readers give the same amount of credence to "not significant" results from small scale studies as they give to significant results from large scale studies (Wood and Lakhani 1997).

### REFERENCE

Wood, J.L.N. and Lakhani, K.H. (1997) Deafness in Dalmatians: Does Sex Matter ? Journal of Veterinary Internal Medicine (submitted).

<sup>1</sup> Centre for Preventive Medicine, Animal Health Trust, Newmarket, Suffolk, CB8 7 DW, England