Letter to the Editor

J Vet Intern Med 2014;28:1380–1381
10.1111/jvim.12417

Dear Editors,

In the article “Interpretation of odds and risk ratios,” O’Connor gives an excellent review of how an odds ratio (OR) can be a biased estimate of a risk ratio (RR), and how the extent of bias is directly related to the baseline risk of the outcome under study. This is the concept behind the “rare disease assumption” whereby the OR provides a good estimate of the RR when the baseline risk is low, but overestimates the effect (protective or harmful) as the baseline risk increases. This causes the OR to be further from 1.0 than the RR.

It is critical to realize that the rare disease assumption does not apply to properly designed case-control studies. Remember that in a case-control study, the investigator determines how common the disease is in the population that enters the study. If the study calls for 1 control per case, the prevalence of disease in the study group is 50%. For 2 controls per case, it is 33%. It makes no difference whether the condition is rare or not in the source population that gives rise to the study population, the study population will typically have 25% to 50% cases.

This fact may lead one to conclude that the OR in a case-control study will always grossly overestimate the RR that one would have calculated if a cohort study had been done in the source population that gave rise to the cases and controls. This conclusion is typically incorrect. In fact, depending on the technique used to sample controls, the OR can provide a completely unbiased estimate of the RR or incidence rate ratio (IRR).

To understand why this is the case, it is helpful to review the concept of a nested case-control study. This is a case-control study performed on members of a cohort study, usually for reasons relating to statistical efficiency or cost. For instance, imagine that you are performing a cohort study to estimate the 12-month risk of congestive heart failure (CHF) in dogs with mitral regurgitation. Initially, you planned to investigate some number of clinical and laboratory findings as risk factors. You enroll 200 dogs, 30 of whom develop CHF over the study period. After the study is over, you decide that it would be interesting to see if a baseline NT-proBNP concentration above 1500 pmol/L was predictive of CHF development. Blood was collected and stored for further testing at the time of enrollment. The RR is calculated as the proportion of the exposed population (NT-proBNP > 1500) who develop disease divided by the proportion of the unexposed (NT-proBNP < 1500) who develop disease.

The following table gives the hypothetical results of the cohort study. There are 60 dogs with high NT-proBNP and 140 dogs in the unexposed group. At the end of the 12-month period, 20 exposed dogs and 10 of the unexposed dogs developed CHF.

	Exposed	Unexposed	Total
Cases	20	10	30
Non cases	40	130	170
Total	60	140	200
Risk	0.3333	0.0714	RR = 4.67
Odds	0.5000	0.0769	OR = 6.50

The OR calculated from cohort data is 6.50, which is an overestimate of the RR.

Let’s add a twist to the scenario. You don’t have enough money to measure NT-proBNP on all 200 dogs enrolled in the study. You have enough to measure the concentration in all 30 cases, but can only measure the concentration in an additional 50 animals. This provides the opportunity to do a nested-case control study. There are a couple of general points about the modern understanding of case-control studies that should be mentioned. First, the purpose of the control group is not to provide a comparison for the case group—the comparison is still between exposed and unexposed, just as in a cohort study. Instead, the control group is meant to represent the exposure distribution of the source population that gives rise to the cases. In a nested case-control study, that source population is simply the population that makes up the cohort. It is useful to realize that every case-control study is nested in a hypothetical cohort that defines the source population your cases came from. Here’s a critical point that seems counter intuitive—it is perfectly acceptable for patients to appear in both the case and control group. In the cohort study, dogs appeared in both the numerator and denominator of the risk calculation. The control group is a substitution for the denominator information, so it makes sense for an individual to appear in the case and control group.

For our study, the control group is a random sample from all 200 dogs. If we take a random sample, the exposure distribution in the sample should, apart from sampling error, match the exposure distribution in the source population. In the cohort, 30% of the dogs had NT-proBNP above the cutoff. Therefore, in our sample of 50 controls, we expect 15 exposed and 35 unexposed. The case group is the same, with 20 exposed cases and 10 unexposed cases. The odds ratio is $(20/15)/(10/35) = 4.67$. It is an unbiased estimate of the risk ratio, despite the fact that the outcome was very common in the source population. What would happen if we restricted our controls to the dogs who did not develop CHF over the 12-month period? There were 170 dogs in this group, of whom 40 had elevated NT-proBNP. Thus, the exposure prevalence in non-cases is 23.5%. If we could study fractions of dogs, our random sample of 50 would have 11.76 exposed and 38.24 unexposed. This gives an OR of 6.5, which is identical to the OR calculated from the cohort data. This OR is a biased estimate of the RR.

In cohort studies with loss to follow-up, the IRR is preferable to the RR because it accounts for unequal follow-up times. A case-control OR will equal the IRR in the source population if controls are sampled with a likelihood proportional to their follow-up time. In this way, the ratio of exposed to unexposed controls will equal the ratio of exposed to unexposed follow-up time. This is the logic behind time-matched case-control studies or density sampling.

In conclusion, as O'Connor concluded, researchers should generally avoid presenting odds ratios in cohort studies or clinical trials as they will overestimate the effect if interpreted as risk or rate ratios. However, if controls are selected correctly, the OR

from a case-control study provides an unbiased estimate of either the RR or IRR. This is true no matter how common the outcome is in the source population that gives rise to your cases and controls. Beware of studies that exclude controls who later became cases. The OR in such a study is not a valid estimate of the RR or IRR in the source population.

Sincerely,

Steve Valeika DVM, Ph D

*Animal Hospital South,
Asheville, NC*