Journal of Veterinary Internal Medicine
Open Access

Standard Article

J Vet Intern Med 2017;31:1035–1042

What's in a Name? The Incorrect Use of Case Series as a Study Design Label in Studies Involving Dogs and Cats

J.M. Sargeant, A.M. O'Connor, J.N. Cullen, K.M. Makielski, and A. Jones-Bitton

Background: Study design labels are used to identify relevant literature to address specific clinical and research questions and to aid in evaluating the evidentiary value of research. Evidence from the human healthcare literature indicates that the label "case series" may be used inconsistently and inappropriately.

Objective: Our primary objective was to determine the proportion of studies in the canine and feline veterinary literature labeled as case series that actually corresponded to descriptive cohort studies, population-based cohort studies, or other study designs. Our secondary objective was to identify the proportion of case series in which potentially inappropriate inferential statements were made.

Design: Descriptive evaluation of published literature.

Participants: One-hundred published studies (from 19 journals) labeled as case series.

Methods: Studies were identified by a structured literature search, with random selection of 100 studies from the relevant citations. Two reviewers independently characterized each study, with disagreements resolved by consensus.

Results: Of the 100 studies, 16 were case series. The remaining studies were descriptive cohort studies (35), population-based cohort studies (36), or other observational or experimental study designs (13). Almost half (48.8%) of the case series or descriptive cohort studies, with no control group and no formal statistical analysis, included inferential statements about the efficacy of treatment or statistical significance of potential risk factors.

Conclusions: Authors, peer-reviewers, and editors should carefully consider the design elements of a study to accurately identify and label the study design. Doing so will facilitate an understanding of the evidentiary value of the results.

Key words: Canine; Case series; Descriptive cohort study; Feline; Population-based cohort study; Study design; Veterinary.

Clinical research provides knowledge that aims to be directly applicable to veterinarians to aid in the diagnosis, treatment, or prevention of disease, or to understand the clinical course or prognosis of a disease. Clinical research encompasses a variety of study designs, each best suited to addressing specific types of research questions. The strongest evidence for efficacy of a treatment that is amenable to randomization is provided by well-conducted systematic reviews and meta-analyses that synthesize results from multiple well-conducted randomized controlled trials (RCT), followed by individual well-conducted RCTs, and then by other designs.[1]

However, not all clinically relevant questions can be addressed with a RCT. Observational studies are particularly valuable for interventions that are not amenable to randomization, where ethical considerations preclude the use of a RCT design, and for evaluating risk factors or exposures, such as age or sex, that cannot be randomized. Although there are a variety of observational study designs[2] (Table 1), the one with the highest evidentiary value for evaluating potential risk factors (or other exposures) usually is considered to be the cohort study. In the exposure-based ("classic") cohort study, outcome-free subjects are purposively selected for the study population based on their exposure status and followed over time to compare the incidence of the outcome of interest among risk factor groups. A variation is the population-based cohort study.[3] In this design, a cohort of individuals who do not have the outcome of interest but who have a particular distinguishing characteristic is selected as the study population. Individuals within the selected cohort then are characterized based on their exposure status, and followed over time to compare the incidence of the outcome among exposure groups.[2] A distinguishing feature of these 2 variations of the cohort design is that there are at least 2 levels of the exposure variable and at least 2 levels of the outcome variable represented in the study population. Thus, these are considered analytical designs and are appropriate for testing hypotheses.

In some instances, the purpose of a study was to solely estimate the incidence risk or incidence rate of an outcome by selecting a cohort of individuals without the outcome of interest and following them over time to calculate the incidence measure. The subjects are enrolled based on an exposure status that defines the cohort. In some instances, eligibility for the cohort may be based on geography or membership in a group (e.g, a herd or a veterinary teaching hospital record system). In other instances, eligibility for the cohort may be based on having a specific disease or condition, and the

From the Centre for Public Health and Zoonoses, University of Guelph, Guelph, ON Canada (Sargeant, Jones-Bitton); Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, ON Canada (Sargeant, Jones-Bitton); Department of Veterinary Diagnostic and Production Animal Medicine, Iowa State University, Ames, IA (O'Connor, Cullen, Makielski).

Corresponding author: J. M. Sargeant, Centre for Public Health and Zoonoses, 50 Stone Road, University of Guelph, Guelph, ON, Canada N1G 2W1; e-mail: sargeanj@uoguelph.ca

Submitted December 22, 2016; Revised March 10, 2017; Accepted April 19, 2017.

DOI: 10.1111/jvim.14741

Table 1. Overview of key features of observational study designs.

Study Design	Selection and Categorization of Study Population	Analysis
Case-control	Based on the presence/absence of an incident (preferred) or prevalent outcome, then categorized into exposure groups	Compare distribution of exposure between outcome groups
Cross-sectional	Sampling without regard to exposure or outcome status, then categorized into groups for exposure status and outcome status	Compare prevalence of the outcome between exposure groups, calculate prevalence of outcome in the population and distribution of the exposure in the population
Exposure-based cohort	Purposive sampling of exposed individuals and nonexposed individuals, all study subjects are outcome negative at the time of sampling and followed over time to determine incidence of an outcome	Compare incidence of outcome between exposure groups
Population-based cohort	Study population selected based on a characteristic of interest (e.g, a disease, condition of interest, geographic location) and categorized into exposure groups, all study subjects are free from the outcome of interest (this could be a consequence of the condition of interest) at the time of sampling and followed over time to determine incidence of the outcome of interest	Compare incidence of outcome between exposure groups, calculate incidence of the outcome and distribution of the exposure in the population
Descriptive cohort	Study population selected based on the presence of a characteristic, such as a disease or condition of interest (the exposure), all study subjects are free from the outcome of interest (a consequence of the condition of interest) at the time of sampling and followed over time to determine incidence of an outcome	Calculate incidence of the outcome for the population
Case series	Study population selected based on the presence of a disease or condition of interest (the outcome)	Description of disease condition, may describe prognosis/survival time for individual study subjects

incidence measure to be estimated is a possible consequence of that condition. An example might be selecting all dogs with a specific type of cancer (the exposure) to form the cohort and measuring mortality risk within 12 months of diagnosis as the outcome. Although the label associated with this single group study type is an area of inconsistency in the literature, it has been proposed that these studies be called descriptive cohort studies.[4] Because descriptive cohort studies can have disease status (such as cases of cancer) as an eligible criterion for the cohort, the design could be misinterpreted as a case series.

Case reports (a single case of a defined disease or condition) or case series (multiple cases) describe a disease or other condition of interest (Table 1). The study subjects are enrolled based on the disease or the condition of interest as the outcome, rather than the exposure status. The purpose generally was to describe the presenting history, clinical presentation and disease progression, and diagnosis for the condition of interest. The prognosis is described as the survival experience of individual animals, rather than as an estimate of incidence for the study population. Often, the condition of interest represents a rare or novel disease, a diagnostic or treatment approach, or an unusual clinical presentation of a more familiar disease.[5] Therefore, case reports and case series are useful for alerting veterinarians about a new or unusual disease condition, and they also may provide ideas for novel diagnostic or therapeutic approaches, or ideas for potential risk factors for a disease condition. The remainder of this article will refer to case series, although the principles discussed herein also apply to case reports.

Case series (and descriptive cohort studies) are not appropriate for testing hypotheses related to treatment efficacy or the statistical significance of potential risk factors, because they do not have a comparison group.[5,6] For this reason, case series rank very low as an evidence source for identifying risk factors, evaluating efficacy of interventions, or assessing the accuracy and usefulness of diagnostic approaches.[1]

Evidence from the human healthcare literature suggests that the term "case series" may be used incorrectly.[4,7] Correct classification is important because the designs are intended to address different questions or, for a specific type of question, provide different strengths of evidence. Also, study design labels often are used in literature searches conducted to identify publications of relevance to a specific research question. It is unknown whether inconsistency in the use of the case series descriptor is an issue in veterinary medicine. Because descriptive cohort studies, population-based cohort studies, and case series all can appear to enroll a single group of diseased animals (whereby this disease in both types of cohort studies is the common characteristic that defines the cohort and not the outcome of

interest), we wished to examine whether investigators accurately differentiate and label these 3 designs. Confusion with exposure-based cohort studies is less likely, because these designs purposefully enroll 2 distinct groups and thus are more readily differentiated.

Therefore, our primary objective was to describe the proportion of studies in dogs and cats labeled as case series that actually represented descriptive cohort studies, population-based cohort studies, or other study designs. Our secondary objective was to identify the proportion of case series that made inappropriate inferential statements on treatment efficacy or risk factor identification.

Methods

Ethical approval for this study was not required, because the "study population" consisted of previously published primary research studies. Studies were identified in MEDLINE via PubMed using the search terms in Table 2. The search was conducted June 1, 2015 without language restrictions and with the date of publication restricted to publication on or after January 1, 2010. The titles and abstracts of the citations identified by the search were screened by the first author to ensure eligibility. To be eligible, the authors had to describe the study as a case series in the title or abstract, and the study population had to be cats or dogs or both. Studies that included other species, in addition to cats or dogs, were excluded.

A random number generator[a] then was used to select 100 relevant citations for inclusion. The full text of the selected studies was obtained using available University of Guelph resources or open access sites. Using the full publication, eligible articles were characterized with a form developed in DistillerSR.[b] A draft of the characterization form was pretested by 3 coauthors (JMS, JNC, AMOC) on a sample of 3 studies selected by JMS to illustrate different types of eligible studies. Based on the results of the pretest, modifications to the wording of the questions were incorporated and the pretest was repeated. The studies used in the pretest were not among the 100 articles included in the results. The final form included 16 questions, 12 of which had preidentified response options, but the form was formatted such that, if a new response was identified, a reviewer could permanently add that response to the form.

The 16 questions included 2 questions to confirm eligibility. If an article was deemed ineligible at this stage, or was not available in English, it was replaced by another randomly selected article from the search. If an article explicitly described both a case series and another design (e.g, "a case series and case-control study"), then only the case series component of the article was used for categorization.

The questions included publication characteristics such as year of publication and species (dog, cat, both). A checklist question was used to describe the type of disease condition by body system (e.g, neurological or respiratory) or by disease type (e.g, cancer). If the disease type was unclear, author KM was consulted for her expertise in small animal veterinary medicine. Additional questions were included to determine the method of selecting the study population (census of animals with a specific disease or condition of interest during a specified time interval; convenience sample; based on a probabilistic sampling scheme; other; not reported) and the method by which cases were selected for inclusion (existing records [retrospective]; sequentially [prospective]; existing records and sequentially; unclear).

A checkbox question was used to collect data on the type of study results reported. The reviewers selected all of the following elements that were included in the results of each study: description of a disease or condition, including signalment, in a single group of animals; comparison of signalment between ≥2 groups; description of a novel diagnostic, therapeutic, or surgical approach in a single group; estimation of the risk or rate of an outcome (including mean or median survival time) in a single group in which there was a longitudinal component (i.e, incidence measure); comparison of the risk or rate of an outcome among groups in which there was a longitudinal component (i.e, comparison of incidence measures); and, other results (including before-after comparison of clinical scores or test results, correlation or agreement among diagnostic tests, and estimates of the prevalence of a condition).

To address the secondary objective, 1 question asked whether the researchers statistically tested ≥1 hypotheses. If no hypotheses were tested, an additional question asked whether the researchers made inferential statements about risk factors or treatment efficacy. If the answer to this question was "yes", a text box was provided for the reviewer to extract the inferential statements.

The remaining questions were in text box form to allow the reviewers to enter the author description of the study design, total sample size, type of study design used if the study description did not correspond to a case series, and general comments. In some instances, 2 sample sizes were reported in a publication (e.g, multiple surgeries described per animal or multiple limbs described per animal). In such cases, the number of animals was recorded.

Two of the 3 reviewers for this project (JMS, JNC, AMOC) independently categorized each article using the structured questions, with any disagreements resolved by consensus. Where consensus could not be reached, a third reviewer was consulted for adjudication.

To determine the actual study design, 2 approaches were used. First, if the reviewers (who all had expertise in epidemiology) determined that the study was an experimental design, a diagnostic test evaluation study, or an observational study other than a population-based cohort study, they recorded the study type. Experimental studies were those in which the investigators allocated individuals to ≥1 treatment or exposure groups. Diagnostic test evaluation studies were studies in which ≥2 tests were compared using measures of agreement or in which a test was compared to a gold standard to calculate sensitivity and specificity. Other observational studies (described in Table 1) could be:

- case-control studies, in which study subjects were selected based on the presence or absence of an outcome of interest and the distribution of the exposure was compared among outcome groups;
- cross-sectional studies, in which individuals were sampled without regard to exposure or outcome status and the prevalence of the outcome was compared among exposure groups; or,
- exposure-based cohort studies, in which an exposure-positive group and an exposure-negative group were purposively selected and followed over time to compare disease incidence.

Table 2. Search terms used to identify case series involving dogs and/or cats and published between 2010 and June 2015.

#4	"Search (#3 and #4 and #5)",525
#3	"Search ((dog or dogs or canine or cat or cats or feline))",486020
#2	"Search (""2010/01/01""[Date - Publication] : ""3000""[Date - Publication])",5211487
#1	"Search ""case series""[Title/Abstract]",42210

For the remaining studies, the question related to the type of included results determined whether each study corresponded to 1 of 3 choices:

- a "true" case series, in which the authors described a disease or condition (presenting history, diagnosis, clinical presentation and progression, or prognosis in individual animals) or described a novel disease, condition, or treatment approach, but did not include a longitudinal calculation of an incidence measure (risk or rate) for the study population or compare the risk of an incidence outcome between groups, or both;
- a descriptive cohort study, in which the authors included a longitudinal calculation of an incidence measure (risk, rate, or survival time) for the study population but did not compare the risk of an incidence outcome among groups; or,
- a population-based cohort study, in which the study group was selected based on having a specific disease or condition and an incidence outcome was compared among ≥2 exposure groups.

The rationale for the sample size was pragmatic, as no prior information on the conduct and reporting of veterinary case series was available on which to base a sample size calculation. A sample size of 100 was chosen because previously conducted studies of similar exploratory style used this sample size.[8–12]

As a post priori step, 1 reviewer (JMS) reevaluated all of the studies that had been categorized as including an estimate of incidence risk or rate to determine whether or not this incidence determination was a stated objective of the publication.

Descriptive statistics were conducted to illustrate the features of all selected studies and to provide descriptive information on characteristics of the sample population for each of the 3 target study design types (i.e, case series, descriptive cohort, population-based cohort). No statistical hypotheses were tested.

Results

The search identified 525 citations. Of these, 81 were not eligible based on screening of the title and abstract, because the authors did not explicitly state that the study was a case series (n = 30) or because the study was not conducted in the target species (n = 51). One of the selected articles was in German; it was excluded and replaced with the next article in the random number sequence.

Descriptive information on the selected studies is presented in Table 3. The majority of case series studies included dogs only (80/100). The year of publication was relatively evenly distributed, given that 2015 was only a partial year because the literature search was conducted in June. The most common conditions studied related to oncology or orthopedic diseases. Nineteen journals were represented, with the number of included studies per journal ranging from 1–48. Two journals comprised 70% of the included publications. Most (60/100) studies were described by the authors as "case series" with the word "retrospective" also included in the design description.

Although all of the included studies were described in the title or abstract as case series, there were 13 studies for which the actual design and conduct of the study corresponded to an experimental design or an observational design other than a case series, descriptive cohort study, or population-based cohort study. Five studies were diagnostic test evaluation studies, 4 were experimental designs (nonrandomized clinical trials, before-after trials, or deliberate disease challenge trials), 2 were cross-sectional studies, and 2 were case-control studies. No exposure-based cohort studies were identified.

Table 3. Descriptive information on 100 studies of dogs and/or cats described by the authors as case series and published between 2010 and June 2015.

	Frequency in Each Category
Species studies	
Dogs only	80
Cats only	12
Dogs and cats	8
Year of publication	
2010	16
2011	18
2012	18
2013	15
2014	24
2015[a]	9
Disease condition or system	
Oncology	20
Orthopedic/skeletal/lameness	19
Neurological	9
Urinary/renal	9
Gastrointestinal	6
Cardiology	5
Respiratory	5
Emergency/critical care	3
Endocrine/metabolic	3
Hepatic	3
Reproductive	3
Toxicology/poisoning	3
Other	12
Author description of study design	
Case series	29
Case series, with "retrospective" included in description	60
Case series, with "prospective" included in description	6
Retrospective and prospective case series	1
Case series plus another study design	4

[a]January to June.

Descriptive statistics for the remaining 87 studies, by study type (case series, descriptive cohort study, or population-based cohort study), are presented in Table 4. Most studies (71/87) were study designs other than case series. Specifically, 35 (of 71) studies included an estimate of incidence risk or rate and had only 1 cohort (i.e, no comparison group), representing a descriptive cohort design. Of these, the estimation of ≥1 incidence outcomes was stated as a study objective for 19 studies (although the word "incidence" was seldom used), estimating risk or rate was not a stated objective for 13 studies, and it was unclear whether or not it was an

Table 4. Descriptive statistics for 87 studies of dogs and/or cats described by their authors as case series, categorized the review authors as case series, descriptive cohorts, or population cohort studies.

	Case Series (n = 16)	Descriptive Cohort (n = 35)	Population Cohort (n = 36)
Selection of study population			
Census of animals with the disease or condition of interest during a specified time interval (i.e, all eligible)	5	25	36
Based on probabilistic sampling scheme	0	0	0
Other	1	2	0
Not reported	10	8	0
Source of study population			
Existing records	6	29	35
Prospective (sequentially)	1	3	0
Both	8	3	1
Unclear	1	0	0
Sample size			
Mean	6.6	23.6	118.1
Standard deviation	4.4	27.2	168.2
Range	6–16	4–113	11–808

objective for 3 studies. Thirty-six of the 71 studies were population-based cohort studies or included a population-based cohort component. Because our study objectives were descriptive, no formal statistical comparisons among study types were conducted. Numerically, most studies represented a census of eligible animals sourced from existing records, although the selection of the study population was not described in a substantive number of studies (n = 18). The case series studies tended to have a smaller sample size, although there was a range of sample sizes for all of the designs.

Of the 51 studies that did not include a parallel comparison group for an exposure variable (16 case series and 35 descriptive cohort studies), 10 (19.6%) formally tested ≥ 1 hypotheses. All 10 were descriptive cohort studies and included hypothesis tests on comparisons of before-after clinical scores within animal, or comparisons between animal signalment and an outcome other than the disease used to define the "case". Of the 41 studies in which no formal tests of statistical significance were conducted, 3/16 case series (18.8%) and 17/25 descriptive cohort studies (68.0%) made unsubstantiated inferences about treatments or risk factors for the outcome. Examples included statements such as, "This study provides evidence that use of [treatment *x*] is efficacious for management of dogs with [disease *y*]" or, "There was no correlation between survival time and tumor size".

Discussion

Our results suggest that most of the studies in the canine and feline veterinary literature that are described as case series studies actually are not case series; the majority are descriptive cohort studies or population-based cohort studies. This finding is consistent with observations in the human healthcare literature.[4,7] The distinction is important, because different designs are used to address different types of questions and their evidentiary value for estimating incidence or for testing hypotheses also differs.

The confusion in terms of labeling the different designs may be related to the use of the word "case" in case series. In a case series, animals are selected for inclusion in the study because they have some disease or condition of interest, the "case".[13] In a descriptive cohort study, individuals also can be selected for inclusion in a study because they have a disease or condition of interest, but that disease or condition defines the cohort, with the outcome being some possible consequence of that exposure (e.g., recovery, complications, death).[4] Similarly, in a population-based cohort study, individuals are selected for the study because they have a characteristic (e.g, disease, condition of interest), which defines the cohort.[2,3] In this instance, they are categorized into ≥2 levels of some additional factor (the "exposure") and the outcome being investigated is a possible consequence of the exposure of interest.

For example, suppose a group of researchers was interested in an unusual neurological disorder as the disease or condition of interest. If researchers identified a group of animals with the disorder and described the signalment and clinical presentation, as well as potentially describing diagnostic approaches and treatment, the study would be a case series. Enrollment was based on the disease condition as the outcome of interest. The study also might report the time to death for the individual animals included in the study population (or note survival time at the time of writing for animals still alive). For instance, if 5 cases were included in the case series, the authors might report that subject 2 died 4 months postdiagnosis, subject 5 died 1 year postdiagnosis, and subjects 1, 3, and 4 were still alive at the time the report was written; this would represent a case series. If, however, the researchers identified a group of animals with the neurological condition (either prospectively or retrospectively) and determined mortality risk (at the study population level), this would mean that the study was a descriptive cohort study. In this instance, the authors might report that the mortality risk for the study population (i.e, at a population level) was 40%, with an appropriate measure of variability, such as a confidence interval. If the researchers identified a group of animals with the neurological condition of interest and categorized the animals based on the severity of neurologic signs at the time of presentation (t0) (or, if retrospective, were able to extract this information from the hospital records), then an incidence outcome, such as mortality risk at some later time (t1), could be compared among the levels of severity at initial presentation. This design has ≥2 levels of the exposure (severity level) and ≥2 levels of the outcome (average survival time or proportion dead at t1), and therefore

it is an appropriate design for hypothesis testing. This scenario describes a population-based cohort study. The example could be extended to describe an exposure-based cohort study, wherein the source population would be animals with the neurological disorder. An outcome-negative group of animals positive for an exposure of interest and an outcome-negative group of nonexposed individuals would be purposively selected, and an incident outcome could be compared among exposure groups. The exposure-based cohort is also an appropriate design for hypothesis testing.

With the case series design, there is only 1 group, the cases (animals with the disease or condition of interest). Evaluating the efficacy of treatments, or identification of risk factors, requires the use of a control group, a group either not treated or treated with an alternative treatment or, in the case of risk factors, groups with and without the putative risk factor. A comparison group is essential to evaluate a hypothesis, because it allows quantification of the role of chance (sampling error) as an explanation for the observed difference in outcome. Additionally, without a control group, it is not possible to determine whether any improvements (or declines) in health are a function of natural disease progression, as opposed to being causally associated with the intervention (treatment). Thus, it is not appropriate to make inferential statements on treatment efficacy or to identify risk factors using a case series design. A rare exception might be a disease with an extremely certain outcome (e.g, the extremely high case fatality associated with clinical rabies), in which any treatment associated with survival might be considered efficacious. Interestingly, although such a case report has been published,[14] the treatment proposed has not resulted in additional treatment successes and is no longer recommended.[15] Case series can provide observations related to possible risk factors or treatments, but such statements should be described as hypotheses, rather than results. These observations then should be evaluated using study designs that involve control/comparison groups and have stronger evidentiary value for addressing questions of efficacy or risk. It is concerning that we identified a relatively high proportion of case series that inappropriately made inferential statements regarding treatment efficacy. Authors of case series should be cautious when wording statements related to treatments or risk factors and likewise readers should be cautious when interpreting them.

The descriptive cohort study is an appropriate design for estimating incidence if the study population is representative of the source population and the source population is representative of the target population, and if a consistent inception point into the study or animal time at risk can be quantified. The validity of incidence estimates therefore is related to the method of sampling the population and the source of the study participants. For instance, if the study population was a census or probabilistic sample of eligible individuals from a teaching hospital database or was based on sequential enrollment of eligible animals in the teaching hospital, then it might be reasonable to conclude that the study population was representative of the source population (relates to internal validity of study). However, if the cases seen in the teaching hospital had a more severe clinical presentation or had different comorbidities than animals in the general population with the condition then, assuming that the general population was the target population, the source population would not be representative of the target population (low external validity). Our finding that information on the selection methods for study subjects was not reported in all studies is of concern. Although not specific to descriptive cohort designs, the STROBE-Vet guidelines recommend that information on the target, source, and study populations be reported; these guidelines also provide information on the level of detail that should be included.[16,17]

A common inception point means that the time of entry of individuals into the cohort is clearly defined, and preferably is the same for all study subjects. This is important because, if the study subjects are observed for variable amounts of time, or if they are at different points in the progression of disease when they enter the study, then it is difficult to interpret incidence estimates such as mortality risk or survival rates.[13] Populations can be closed or open, and the distinction is important when deciding which incidence measure to use and when interpreting incidence estimates.[18] Closed populations are those in which individuals eligible for enrollment are defined by an event (e.g, birth year, type of surgery) that precludes other individuals from entering the study population. In an open population, individuals can enter or leave the population, meaning that time at risk during a study can differ among individuals. In some instances, an open population can "become closed" by considering time from a specific event, rather than using calendar time. For example, in a descriptive cohort study, the study population may be open if individuals are selected with different amounts of time after acquiring the disease or condition that defines the cohort or if the follow-up period is variable. In these cases, it may be difficult (or inappropriate) to estimate incidence risk, although incidence rate could be calculated. However, if entry into the cohort is defined as a specific event (i.e, the open population becomes closed), such as the time of surgery, and the follow-up time is consistent among study subjects, then valid estimates of incidence risk can be derived. In our study, approximately half of the descriptive cohort studies included an estimate of an incidence outcome as a specific objective, and evaluating whether the study populations were open or closed was beyond the scope of our study. In a case series, the authors may comment on the outcomes, such as the number alive at the time the study period ended, but the time at risk for individual animals is not known or the duration of follow-up may vary among animals. In our study, it was difficult in some publications to determine whether individual time at risk was known or consistent. Therefore, some of the studies categorized as descriptive cohort studies may actually have been case series, but with an estimate of an

incidence outcome that did not include a consistent or known time at risk.

Sample size is not a distinguishing feature for case series versus descriptive cohort studies and there can be large case series or small population cohort studies, as observed here. However, the usefulness of incidence estimates in descriptive cohort studies varies with sample size and small sample sizes could lead to estimates with such low precision that they are poorly informative. For example, consider 2 descriptive cohort studies. In the first study, there are 4 dogs with the condition of interest and 2 of these dogs die between t0 and t1. In the second study, there are 100 dogs of which 50 die between t0 and t1. In both cases, the mortality risk is 50%. However, using standard confidence interval calculations (http://www.openepi.com/Proportion/Proportion.htm), the 95% confidence interval around that mortality risk would be from 1% to 99% for the first study and from 40.2% to 59.8% for the second. In addition, it would be difficult to judge the representativeness of the study population to the source population (and the target population) in the previous example in which only 4 animals were included in the study population. Thus, the usefulness of the mortality risk estimate for any practical purposes is related to sample size.

Finally, 35 of the studies evaluated were actually population-based cohort studies, wherein the criteria for inclusion was a characteristic (disease or condition), study subjects were characterized by status of an exposure variable, and an incident outcome was compared among exposure levels. For exposure factors that are not amenable to randomization, this design provides the lowest risk of bias for estimating the association (and thus has the highest evidentiary value). Labeling these studies as "case series" inappropriately downgrades the level of evidence. This speaks to the need for readers, reviewers, and editors of the scientific literature to be conversant with the key elements of study design as they relate to study validity and risk of bias, rather than relying on the design labels applied by the authors. It also highlights the importance of authors accurately labeling the studies and reporting key design elements.

Limitations

The results of our study were based on 100 studies described by the authors as case series in the canine and feline veterinary literature, which were identified from a single electronic database. We did not identify an *a priori* sample size, because we were not testing an *a priori* hypothesis and we were not attempting to estimate a parameter within an *a priori* stated allowable error. We believe that the sample size was sufficient to illustrate our key findings. It is possible the studies we selected are not representative of the published veterinary literature. We attempted to decrease this risk by randomly selecting studies from the citations identified by the search. However, most studies were published in 1 of 2 journals. Thus, the editorial policies and reviewer characteristics of those journals may have high potential to influence our results. Nonetheless, our search did not target any specific journal and therefore the search results were likely reflective of the proportion of case series by journal among the journals indexed in MEDLINE. Some of the categorizations required judgments and, in some cases, it was difficult to determine some aspects based on the detail reported in the study. To decrease the risk of subjectivity in our results, 2 reviewers independently characterized each study, with any disagreements resolved by consensus.

In summary, our results suggest that the label "case series" often is applied inappropriately to studies of dogs and cats in the veterinary literature. This could have implications when interpreting the evidentiary value of the results. It is important to consider aspects of study design when designating design labels and to report all elements of the design that are required to allow the reader to evaluate the risk of bias and, therefore the evidentiary value of the study results.

Funding

No external funding was obtained for this study.

Conflict of Interest Declaration: Annette O'Connor serves as Associate Editor for the Journal of Veterinary Internal Medicine. She was not involved in review of this manuscript.

Off-label Antimicrobial Declaration: Authors declare no off-label use of antimicrobials.

Footnotes

[a] Microsoft Excel® for MAC, Version 15.27: 161010
[b] Evidence Partners, Ottawa, Canada

References

1. Sargeant JM, Kelton DF, O'Connor AM. Study designs and systematic review of interventions: Building evidence across study designs. Zoon Public Health 2014;61:10–17.

2. Dohoo IR, Martin SW, Stryhn H. Veterinary Epidemiologic Research, 2nd edn. Charlottetown, Prince Edward Island, Canada: VER, Incorporated; 2009.

3. Szklo M. Population-based cohort studies. Epidemiol Rev 1998;20:81–90.

4. Esene IN, Ngu J, El Zoghby M, et al. Case series and descriptive cohort studies in neurosurgery: The confusion and solution. Childs Nerv Syst 2014;30:1321–1332.

5. Kempen JH. Appropriate use and reporting of uncontrolled case series in the medical literature. Am J Ophthalmol 2011;151:7–10.

6. Kooistra B, Dijkman B, Einhorn TA, et al. How to design a good case series. J Bone Joint Surg Am 2009;91(Suppl 3):21–26.

7. Dekkers OM, Egger M, Altman DG, et al. Distinguishing case series from cohort studies. Ann Intern Med 2012;156:37–40.

8. Sargeant JM, O'Connor AM, Renter DG, et al. Reporting of methodological features in observational studies of pre-harvest food safety. Prev Vet Med 2011;98:88–98.

9. Sargeant JM, Thompson A, Valcour J, et al. Quality of reporting of small animal clinical trials and associations with treatment effects. J Vet Int Med 2010;24:44–50.

10. Sargeant JM, Elgie R, Valcour J, et al. Methodological quality and completeness of reporting in clinical trials conducted in livestock species. Prev Vet Med 2009;91:107–115.

11. Sargeant JM, Saint-Onge J, Valcour J, et al. Quality of reporting in clinical trials of pre-harvest food safety interventions and associations with treatment effect. Foodborne Path Dis 2009;6:989–999.

12. Cullen JN, Sargeant JM, Makielski KM, et al. The case-control design in veterinary sciences: A survey. Prev Vet Med 2016;134:179–187.

13. Holmes M, Cockcroft P. Handbook of Veterinary Clinical Research. Oxford, UK: Blackwell Publishing Ltd. 2008.

14. Willoughby RE Jr, Tieves KS, Hoffman GM, et al. Survival after treatment of rabies with induction of coma. N Engl J Med 2005;16:2508–2514.

15. Jackson AC. Human rabies: A 2016 update. Curr Infect Dis Rep 2016;18:38.

16. Sargeant JM, O'Connor AM, Dohoo IR, et al. Methods and processes of developing the Strengthening the Reporting of Observational Studies in Epidemiology – Veterinary (STROBE-Vet) statement. J Vet Intern Med 2016;30:1887–1895.

17. O'Connor AM, Sargeant JM, Dohoo IR, et al. Explanation and elaboration document for the STROBE-vet statement: Strengthening the reporting of observational studies in epidemiology – veterinary extension. J Vet Intern Med 2016;30:1896–1928.

18. Vandenbroucke J, Pearce N. Point: Incident exposures, prevalent exposures, and causal inference: Does limiting studies to persons who are followed from first exposure onward damage epidemiology? Am J Epidemiol 2015;182:826–833.