ORIGINAL ARTICLE

Explanation and Elaboration Document for the STROBE-Vet Statement: Strengthening the Reporting of Observational Studies in Epidemiology – Veterinary Extension

A. M. O'Connor[1], J. M. Sargeant[2,3], I. R. Dohoo[4], H. N. Erb[5], M. Cevallos[6], M. Egger[6], A. K. Ersbøll[7], S. W. Martin[3], L. R. Nielsen[8], D. L. Pearl[3], D. U. Pfeiffer[9], J. Sanchez[10], M. E. Torrence[11], H. Vigre[12], C. Waldner[13] and M. P. Ward[14]

[1] Department of Veterinary Diagnostic and Production Animal Medicine, Iowa State University, Ames, IA, USA
[2] Centre for Public Health and Zoonoses, University of Guelph, Guelph, ON, Canada
[3] Department of Population Medicine, Ontario Veterinary College, Guelph, ON, Canada
[4] Centre for Veterinary Epidemiological Research, University of Prince Edward Island, Charlottetown, PEI, Canada
[5] Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, NY, USA
[6] Institute of Social and Preventive Medicine, University of Bern, Bern,Switzerland
[7] National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark
[8] Section for Animal Welfare and Disease Control, University of Copenhagen, Copenhagen, Denmark
[9] Department of Production and Population Health, Royal Veterinary College, London, UK
[10] Department of Health Management, University of Prince Edward Island, Charlottetown, PEI, Canada
[11] Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, MD, USA
[12] National Food Institute, Technical University of Denmark, Lyngby, Denmark
[13] Department of Large Animal Clinical Sciences, Western College of Veterinary Medicine, University of Saskatchewan, Saskatoon, SK, Canada
[14] Faculty of Veterinary Science, The University of Sydney, Sydney, NSW,Australia

Impacts

- Observational studies can inform many decisions in companion animal care, livestock production and food safety
- Comprehensive reporting is critical for enabling readers to understanding the validity of studies
- Specific recommendations are need for observational studies to maximize the value of research results.

Keywords:
Case control study; cohort study; cross-sectional study; observational study; reporting guidelines

Correspondence:
A. M. O'Connor. Department of Veterinary Diagnostic and Production Animal Medicine, Iowa State University, Ames, IA, USA.
Tel.: +1 515 294 5012;
Fax: +1 515 294 0172;
E-mail: oconnor@iastate.edu

Received for publication January 10, 2016

doi: 10.1111/zph.12315

Summary

The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement was first published in 2007 and again in 2014. The purpose of the original STROBE was to provide guidance for authors, reviewers and editors to improve the comprehensiveness of reporting; however, STROBE has a unique focus on observational studies. Although much of the guidance provided by the original STROBE document is directly applicable, it was deemed useful to map those statements to veterinary concepts, provide veterinary examples and highlight unique aspects of reporting in veterinary observational studies. Here, we present the examples and explanations for the checklist items included in the STROBE-Vet Statement. Thus, this is a companion document to the STROBE-Vet Statement Methods and process document, which describes the checklist and how it was developed.

Introduction

In veterinary research, observational studies are commonly used to describe the natural history of disease, assess aetiology, and identify and investigate the effect of risk factors. To maximize the value of observational studies, it is critical that they are reported in a manner that facilitates internal and external validity assessment. Reporting guidelines allow researchers to appraise the published findings and potentially apply them to future research or decision-making.

Initially used for intervention (clinical trial) assessments, the CONSORT (Moher et al., 2001; Schulz et al., 2010) and REFLECT statements (O'Connor et al., 2010; Sargeant et al., 2010) were developed to create an experimental and reporting framework for randomized controlled trials and to help authors, reviewers and editors address concerns about incomplete reporting. The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement, first published in 2007 and again in 2014 (von Elm et al., 2007; Vandenbroucke et al., 2007, 2014), provided a similar framework for observational studies. In this document, we provide the rationale behind the revision of STROBE for use in veterinary research and examples of data reporting under the revised guidelines. While much of the STROBE material is directly relevant to veterinary studies, animal health investigations have sufficient unique features to warrant publishing a set of veterinary-investigator-specific guidelines (reference Process document). For example, multiple levels of organization are common in animal populations, and observational studies should account for this when reporting results. Given the importance of population structures when interpreting results, this issue features prominently in the STROBE-Vet extension.

Omission or unclear reporting of important details is a common problem in all types of research reports. Some omissions can seriously limit the utility of the research by either hiding limitations or creating unwarranted doubt about the study conclusions These omissions, in turn, increase research wastage (Fanelli, 2013; Al-Shahi Salman et al., 2014; Chan et al., 2014; Ioannidis et al., 2014; Macleod et al., 2014; Moher et al., 2016). Study results are usually used by people other than the manuscript authors to make decisions. Hence, these users need as much information as possible to judge the validity of the results. Reporting guidelines are designed to reduce critical omissions by providing a checklist of important items to include in the report. Checklists improve author, editor and reviewer compliance with respect to what information should be included in a comprehensive report, making them valuable research-reporting tools (Haynes et al., 2015; Kim et al., 2015).

How to Use this Document

Each item is presented in the same manner: first the item number (1 to 22) with subdivisions and a description of the item, followed by examples that illustrate the reporting approach for the item and a discussion of the rationale for their inclusion. Ideally, the examples chosen would illustrate all of the key concepts, and only those concepts. However, it was not always possible to identify such specific real-world examples from the veterinary literature. The working group decided not to use human healthcare or hypothetical examples. As a consequence, the examples sometimes include additional examples or several examples were needed to illustrate the key concepts. When the explanation for an item was the same as that reported in the original STROBE publication, we used the material *ad verbatim*, with permission from the original authors. Examples of poorly reported items were not included due to space considerations and the consensus that their inclusion would not substantially increase understanding or adoption of the guidelines. A table with the STROBE-Vet checklist is included at the end of this document (Table 1).

Title and Abstract

The purpose of the abstract and title is to quickly allow the reader to identify the topic of the research, the general design of the study, the main results and the implications of the findings.

1 (a). Indicate that the study was an observational study and, if applicable, use a common study design term

Example 1
Title: 'An observational study with long-term follow-up of canine cognitive dysfunction: Clinical characteristics, survival and risk factors' (Fast et al., 2013).
Example 2
Title: 'Case-control study of risk factors associated with *Brucella melitensis* on goat farms in Peninsular Malaysia' (Bamaiyi et al., 2014).

Explanation
Including the study design term in the title or abstract when a standard study design is used, or at least identifying that a study is observational, allows the reader to easily identify the design and helps to ensure that articles are correctly indexed in electronic databases (Benson and Hartz, 2000). In STROBE, item 1a only requests that a common study design term be used. However, in veterinary research, not all observational studies are easily categorized into cohort, case–control, or cross-sectional study designs. Therefore, we recommend including that the study was observational and, if possible, the study design or important design characteristics, for example longitudinal, in the title.

1 (b). Indicate why the study was conducted, the approach, the results, the limitations and the relevance of the findings

Example
Methicillin-resistant *Staphylococcus pseudintermedius* (MRSP) has emerged as a highly drug-resistant small animal veterinary pathogen. Although often isolated from outpatients in veterinary clinics, there is concern that

Table 1. Strengthening the Reporting of Observational Studies in Epidemiology statement checklist for Veterinary medicine (the STROBE-Vet statement)

	Item	
Title and Abstract	1	(a) Indicate that the study was an observational study and, if applicable, use a common study design term (b) Indicate why the study was conducted, the design, the results, the limitations and the relevance of the findings
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	(a) State-specific objectives, including any primary or secondary pre-specified hypotheses or their absence (b) Ensure that the level of organization[a] is clear for each objective and hypothesis
Study design	4	Present key elements of study design early in the paper
Setting	5	(a) Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up and data collection (b) If applicable, include information at each level of organization
Participants [b]	6	Describe the eligibility criteria for the owners/managers and for the animals, at each relevant level of organization Describe the sources and methods of selection for the owners/managers and for the animals, at each relevant level of organization Describe the method of follow-up (d) For matched studies, describe matching criteria and the number of matched individuals per subject (e.g. number of controls per case)
Variables	7	(a) Clearly define all outcomes, exposures, predictors, potential confounders and effect modifiers. If applicable, give diagnostic criteria (b) Describe the level of organization at which each variable was measured (c) For hypothesis-driven studies, the putative causal structure among variables should be described (a diagram is strongly encouraged)
Data sources/ measurement	8[c]	(a) For each variable of interest, give sources of data and details of methods of assessment (measurement). If applicable, describe comparability of assessment methods among groups and over time (b) If a questionnaire was used to collect data, describe its development, validation and administration (c) Describe whether or not individuals involved in data collection were blinded, when applicable (d) Describe any efforts to assess the accuracy of the data (including methods used for 'data cleaning' in primary research, or methods used for validating secondary data)
Bias	9	Describe any efforts to address potential sources of bias due to confounding, selection or information bias
Study size	10	(a) Describe how the study size was arrived at for each relevant level of organization (b) Describe how non-independence of measurements was incorporated into sample size considerations, if applicable (c) If a formal sample size calculation was used, describe the parameters, assumptions and methods that were used, including a justification for the effect size selected
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why
Statistical methods	12	(a) Describe all statistical methods for each objective, at a level of detail sufficient for a knowledgeable reader to replicate the methods. Include a description of the approaches to variable selection, control of confounding and methods used to control for non-independence of observations (b) Describe the rationale for examining subgroups and interactions and the methods used (c) Explain how missing data were addressed (d) If applicable, describe the analytical approach to loss to follow-up, matching, complex sampling and multiplicity of analyses (e) Describe any methods used to assess the robustness of the analyses (e.g. sensitivity analyses or quantitative bias assessment)
Participants	13[c]	(a) Report the numbers of owners/managers and animals at each stage of study and at each relevant level of organization – for example, numbers eligible, included in the study, completing follow-up and analysed (b) Give reasons for non-participation at each stage and at each relevant level of organization (c) Consider use of a flow diagram and/or a diagram of the organizational structure
Descriptive data on exposures and potential confounders	14[c]	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders by group and level of organization, if applicable (b) Indicate number of participants with missing data for each variable of interest and at all relevant levels of organization (c) Summarize follow-up time (e.g. average and total amount), if appropriate to the study design

Table 1. (Continued)

	Item	
Outcome data	15[c]	(a) Report outcomes as appropriate for the study design and summarize at all relevant levels of organization (b) For proportions and rates, report the numerator and denominator (c) For continuous outcomes, report the number of observations and a measure of variability
Main results	16	(a) Give unadjusted estimates and, if applicable, adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders and interactions were adjusted. Report all relevant parameters that were part of the model (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done, such as sensitivity/robustness analysis and analysis of subgroups
Key results	18	Summarize key results with reference to study objectives
Strengths and Limitations	19	Discuss strengths and limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies and other relevant evidence
Generalizability	21	Discuss the generalizability (external validity) of the study results
Transparency	22	(a) *Funding* – Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based (b) *Conflict of interests* – Describe any conflict of interests, or lack thereof, for each author (c) *Describe the authors' roles* – Provision of an authors' declaration of transparency is recommended (d) *Ethical approval* – Include information on ethical approval for use of animal and human subjects (e) *Quality standards* – Describe any quality standards used in the conduct of the research

[a]Level of organization recognizes that observational studies in veterinary research often deal with repeated measures (within an animal or herd) or animals that are maintained in groups (such as pens and herds); thus, the observations are not statistically independent. This non-independence has profound implications for the design, analysis and results of these studies.
[b]The word 'participant' is used in the STROBE statement. However, for the veterinary version, it is understood that 'participant' should be addressed for both the animal owner/manager and for the animals themselves.
[c]Give such information separately for cases and controls in case–control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

MRSP follows a veterinary-hospital associated epidemiology. This study's objective was to identify risk factors for MRSP infections in dogs and cats in Germany. Clinical isolates of MRSP cases (n = 150) and methicillin-susceptible *S. pseudintermedius* (MSSP) controls (n = 133) and their corresponding host signalment and medical data covering the six months prior to staphylococcal isolation were analysed by multivariable logistic regression. The identity of all MRSP isolates was confirmed through demonstration of *S. intermedius*-group specific *nuc* and *mecA*. In the final model, cats (compared to dogs, OR 18.5, 95% CI 1.8–188.0, P = 0.01), animals that had been hospitalised (OR 104.4, 95% CI 21.3–511.6, P < 0.001), or visited veterinary clinics more frequently (>10 visits OR 7.3, 95% CI 1.0–52.6, P = 0.049) and those that had received topical ear medication (OR 5.1, 95% CI 1.8–14.9, P = 0.003) or glucocorticoids (OR 22.5, 95% CI 7.0–72.6, P < 0.001) were at higher risk of MRSP infection, whereas *S. pseudintermedius* isolates from ears were more likely to belong to the MSSP group (OR 0.09, 95% CI 0.03–0.34, P < 0.001). These results indicate an association of MRSP infection with veterinary clinic/hospital settings and possibly with chronic skin disease. There was an unexpected lack of association between MRSP and antimicrobial therapy; this requires further investigation ..... (Lehner et al., 2014).

Explanation

The abstract provides key information that enables readers to understand the key aspects of the study and decide whether to read the article. In STROBE, item 1b recommended that authors provide an informative and balanced summary of what experiments were done, what results were found and the implications of the findings in the abstract. In STROBE-Vet, this item was modified to provide more guidance on the key components that should be addressed. The study design should be stated; however, if the study does not correspond to a named study design such as case–control, cross-sectional and cohort study, then the author should describe the key elements of the study design such as incident versus prevalent cases, and whether or not the selection was based on outcome status (Pearce, 2012). The abstract should succinctly describe the study objectives, including the primary objective and primary outcome, the exposure(s) of interest, relevant population information such as species and the purpose (or uses) of the animals, the study location and dates, and the number of study units. In addition, including the organizational level at

which the outcome was measured (e.g. herd, pen or individual) is recommended. The presented results should include summary outcome measures (e.g. frequency or appropriate descriptor of central tendency such as mean or median) and, if relevant, a clear description of the association direction along with accompanying association measures (e.g. odds ratio) and measures of precision (e.g. 95% confidence interval) rather than *P*-value alone. We discourage stating that an exposure is or is not significantly associated with an outcome without appropriate statistical measures. Finally, because many veterinary observational studies evaluate multiple potential risk factors, the abstract should provide the number of exposure–outcome associations tested to alert the end-user to potential type I error in the study. When multiple outcomes are observed, provide the reader with a rationale for the outcomes presented in the abstract, for example only statistically significant results or the outcome of the primary hypothesis is presented.

Introduction

The aim of the introduction is to allow the reader to understand the study's context and the results' potential to contribute to current knowledge.

2. Background/rationale: explain the scientific background and rationale for the investigation being reported

Example
The syndesmochorial placenta of cattle prevents the bovine fetus from receiving immunoglobulins in utero; therefore, calves are born essentially agammaglobulinemic [].[1] Calves acquire passive immunity by consuming colostrum in the first 24 to 36 h of life []. Inadequate colostrum consumption leads to failure of passive transfer (FPT), which has detrimental effects on calf health and survival. As many as 40% of dairy calves experience FPT []. However, beef and dairy calf management is considerably different, as beef calves generally remain with the cow post-calving and nurse ad libitum, while dairy producers often separate calves from their dams and then provide the colostrum. Hence, the prevalence of and risk factors for FPT in beef calves may vary substantially from those in reports describing dairy calves. . .' (Waldner and Rosengren, 2009).

Explanation
The scientific background provides important context for readers. It describes the focus and gives an overview of what is known on a topic and what gaps in current knowledge are addressed by the study. Background material should note recent pertinent studies and any reviews of pertinent studies. The background section should also include the anticipated impact of the work.

3 (a). Objectives: state-specific objectives, including any primary or secondary pre-specified hypotheses or their absence

Example
The objective of this study was to investigate the effect of track way distance and cover on the probability for lameness in Danish dairy herds using grazing. We hypothesised that short track distances with added cover would be associated with the lowest lameness prevalence (Burow et al., 2014).

Explanation
Objectives are the detailed aims of the study. Well-crafted objectives specify populations, exposures and outcomes, and parameters that will be estimated. They might be formulated as specific hypotheses or as questions that the study was designed to address. In some situations, objectives might be less specific, for example in early discovery phases. Regardless, the report should clearly reflect the investigators' original intentions.

3 (b). Ensure that the level of organization is clear for each objective and hypothesis

Example
There were three objectives for this study: (1) to quantify the standing and lying behavior, with particular emphasis on post-milking standing time, of dairy cows milked 3 × /d, (2) to determine the cow- and herd-level factors associated with lying behavior, and (3) to relate these findings to the risk of experiencing an elevation in somatic cell count (SCC) (Watters et al., 2013).

Explanation
A full explanation is provided in Box 4: Organization structures in animal populations

Methods

The aim of the methods section is to describe what experiments were planned and performed in sufficient detail for the reader to understand them; judge whether they were adequate with respect to providing reliable, valid answers to the objectives and hypotheses; and assess whether deviations from the original research plan were justified.

[1]In examples quoted directly from the literature, citations were removed and replaced with [] to indicate that the statement was referenced in the original document.

Box 4

Organization structures in animal populations

Many animal populations occur in organizational structures, which results in individual animals (or groups of animals) not being independent from one another (Schukken et al., 2003; Sargeant et al., 2010). These organizational structures might be hierarchical, such as those related to housing (animals within barns, barns within farms, farms within production systems, production systems within regions) or genetics (piglets within sows, calves within dams, daughters within sires). Animal populations can also be non-independent but not hierarchical. For example, beef calves from several cow-calf farms might be transported to multiple feedlots, where calves from multiple farms commingle in pens. Calves from the same farm or housed in the same pen or feedlot probably have more exposures in common than calves at a different farm or in a different pen or feedlot. Such organizational structures imply non-independence, which will influence the actual number of observational units in the study and power in the statistical analyses. Therefore, the non-independence must be accounted for in the study design or adjusted for in the data analysis (European Food Safety Authority, 2013).

Further, the study's end-users might be interested in different hierarchy levels. Thus, it is essential that the authors clearly state what level is being studied. For example, for a particular disease, producers and veterinarians might focus on the disease prevalence within herds and factors associated with individual risk of developing disease (Rose, 2001). However, company officials might be interested in the prevalence of positive herds within a production system and factors associated with a herd being positive or with high or low prevalence (Rose, 2001). Government officials might concentrate on differences in the prevalence of positive herds across regions of a country or among countries. It is also possible to report the outcomes of interest at different organizational structure levels in a single study (Kadohira et al., 1997; McDermott et al., 1997; De Vliegher et al., 2004). Given this complexity, authors must ensure that readers are aware of the organizational level(s) that exist within the study population and the level at which variables are measured and summarized. This information allows the reader to (i) decide if the paper is of interest and (ii) assess experimental approaches for biases, which might differ based on the hierarchy level summarized. A diagram showing the organizational structure might be helpful to convey this information.

The organizational structure is relevant to numerous parts of a publication. In particular, we advise providing information about the study population's organizational structure in items 3, 6, 7, 12, 13, 14 and 15. Here, we provide two study examples along with a description of how to report organization structures in items 3, 6, 7 and 12.

Example 1. A hypothetical multiclinic study of demographic factors affecting survival of dogs with osteosarcoma

Item 3 would describe the study objective: to understand demographic factors that impact a dog's survival time. For this theoretical example, the hypothesis is that dog age is associated with reduced survival time in individual dogs.

Item 5 would describe the clinics and clinicians participating in the study and indicate that they are a likely source of non-independence.

Item 6 would describe the eligibility criteria for selecting clinics, clinicians and clients and dogs for the study.

Item 7 would define the outcome and other variables, as well as the organizational level for each variable. For this hypothetical example, the measurement level for the outcome was at the individual level such as a dog's survival time. The exposure factors of interest were also at the individual level such as the dog age, dog weight and dog breed.

Item 8 would describe how each of the variables listed in item 7 was measured and state that all of these measurements were performed at the individual level (the dog level).

Item 12 would describe how the analysis approach accounted for the impact of the organization structure such as dog non-independence, nested within clinics and clinicians.Example 2. A hypothetical multifarm study of factors affecting the prevalence of *Salmonella* in swine barns

Item 3 would describe the study objective: to understand barn-, site- and company-level characteristics associated with the prevalence of *Salmonella* in swine barns. In the example study, the hypothesis was that the prevalence of *Salmonella* is higher in barns where birds are observed.

Item 5 would state that the pigs are nested within barns, the barns are nested within sites, and the sites within companies. Other possible sources of non-independence (e.g. if farms are nested geographically) should also be stated.

Item 6 would describe the criteria for selecting the companies, the sites within each company, the barns within each site and the pigs within each barn. It would include at what organizational levels, convenience sampling was used. For example, in our hypothetical study, researchers used a relationship with a production company to gain access to a production site. They also used convenience to decide which production sites to study and selected all barns on each site to be surveyed. Then, they randomly selected 30 pigs within each barn to obtain barn-level estimates of *Salmonella* prevalence.

Item 7 would define the outcome and other variables, ensuring the organizational level is stated clearly. For this study, the outcome of interest was the prevalence of *Salmonella* in each barn. The exposure variables of interest were feed type (a site-level variable), and potential confounders included the feed mill used (a site-level variable) and the presence of birds in barns (a barn-level variable).

Item 8 would describe how each of the variables defined in item 7 were measured. In this study, it would be important to clarify that *Salmonella* status was measured in pigs (an individual-level variable) and the prevalence was summarized as a proportion, so it could be expressed as a barn-level outcome. The laboratory approach for determining *Salmonella* status should be described here. This item should state that data regarding the feed mills used at each site were obtained from company records and the presence of birds was determined using a questionnaire administered to the site manager. The validity of that questionnaire should be described as a component of this item.

Item 12 would describe how the analysis approach accounted for the organizational structure such as non-independence of barns within farms, farms within regions, farms within the production system. Because the outcome was measured at the barn level (as clarified in item 7), authors would need to account for the clustering of pigs within barns.

4. Study design: present key elements of study design early in the paper

Example

A cohort study was performed on two farrow-to-finish farms (A and B) in two farrowing rooms (cohorts) per farm. Sows were examined for the presence of *A. pleuropnemoniae* infection by collection of blood and tonsil brush samples approximately three weeks before parturition. The proportions of colonization at litter and individual piglet level were determined three days before weaning and associations with dam parity and sow serum and brush sample results were evaluated (Tobias et al., 2014).

Explanation

We advise presenting key elements of study design early in the methods section (or at the end of the introduction) so that readers can understand the basics of the study. For example, if the authors used a cohort study design, which followed animals or animal groups over a particular time period, they should describe the group that comprised the cohort and their exposure status. Similarly, if the investigation used a case–control design, the cases and controls and their source population(s) should be described.

If a study is a variant of the three main study types (cohort, case–control or cross-sectional), there is an additional need for clarity. Authors can provide a clear description of the study design by including the following key elements: (i) the timing of study population enrolment with respect to the occurrence of the outcome such as after or prior to, (ii) the role of exposure status on enrolment such as enrolled based on exposure or not, (iii) the role of outcome status on enrolment such as enrolled based on outcome or not, (iv) the timing of outcome and exposure determination such as outcome determined before, after or concurrent to exposure determination and (v) if the outcome is a disease, condition or behaviour, whether the outcome represents incidence or prevalence. If the study only estimates prevalence or incidence in a single group, then the authors need to clarify whether the outcome represents incidence or prevalence. This item is intended to give the reader a general idea of the study design. The design specifics are described in detail in subsequent items.

We recommend that authors refrain from calling a study 'prospective' or 'retrospective' because these terms are ill defined (Vandenbroucke, 1991). One usage sees cohort and prospective as synonymous and reserves the word retrospective for case–control studies. A second usage distinguishes prospective and retrospective cohort studies according to the timing of data collection relative to when the idea for the study was developed (Miettinen, 1985). A third usage distinguishes prospective and retrospective case–control studies depending on whether the data about the exposure of interest existed when cases were selected (Rothman and Greenland, 1998).

In STROBE-Vet, we do not use the words prospective and retrospective, nor alternatives such as concurrent and historical. We recommend that whenever authors use these

words, they define what they mean. Most importantly, we recommend that authors describe exactly how and when data collection took place.

5 (a). Setting: describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up and data collection

5 (b). If applicable, include information at each level of organization

Example

This study was conducted in Afar and Tigray regions in north-eastern Ethiopia. Two administrative zones (Zone-1 and Zone-4) out of five zones of Afar region were included in the study, and then one district from each zone was selected (Asiyta and Yallo, respectively). Asayita district was selected to include an agro-pastoral production system where irrigation farming is widely prevalent. ... Yallo was selected for its location interfacing with the highland agro-climate in Alamata and Raya Azebo districts where the livestock are moved for grazing and watering during dry season [ ]. There were two distinct agro-ecological climates prevailing in the Afar study area: lowland (<1500 m) and highland (>2300 m). ...

A cross-sectional study was carried out between October 2011 and February 2012 to assess epidemiological factors associated with observed [lumpy skin disease] in the previous two years (September 2009 to October 2011). Three to four Kebeles (the lowest administrative unit next to district in order of hierarchy in Ethiopia) were selected randomly from each district, and 20-30 herds were randomly selected from each Kebele. Herd-owners were selected based on willingness to complete the questionnaire (Hailu et al., 2014).

Explanation

Readers must understand the clinical, demographic, managerial, geographic and temporal contexts in which the study was conducted, so readers will be able to determine the populations to which the study's inferences can be applied. Data from research herds or kennels might not extrapolate to commercial or home settings. Dates are required to understand the historical context of the research, because medical, sociological and agricultural practices can change over time, which, in turn, can affect the prevalence of risk factors, potential confounders, diseases and study methods. Knowing when a study took place and over what period participants were recruited and followed places the study in historical context and is important for the interpretation of results.

6. Participants

6 (a). Describe the eligibility criteria for the owners / managers and for the animals, at each relevant level of organization

Example

Counties were chosen based on the proportion of registered backyard flock owners and location of commercial industries and auction markets. In May 2011, the Maryland Department of Agriculture (MDA) confidentially mailed 1,000 informational letters and return postcards to poultry owners enrolled in the Maryland Poultry Registration Program. Participants were eligible for the study if they lived in Maryland, owned domesticated fowl, and maintained a flock size fewer than 1,000 birds (Madsen et al., 2013).

Explanation

Eligibility criteria might be presented as inclusion and exclusion criteria, although this distinction is not always necessary or useful. Regardless, we advise authors to report all eligibility criteria and also to describe the group from which the study population was selected (e.g. the general population of a region or country), and the method of recruitment (e.g. referral or self-selection through advertisements). Authors of studies involving animal populations should describe the eligibility criteria at all organizational levels (e.g. farm, pen, stable or clinic) for the animals included, and for smaller units within included animals, such as limbs or mammary quarters, if applicable (see Box 4: Organization structures in animal populations).

6 (b). Describe the sources and methods of selection for the owners/ managers and for the animals, at each relevant level of organization

Example

All MRSP isolates identified between October 2010 and October 2011 inclusive were considered. MSSP isolates were selected throughout the study period using simple randomization on www.randomizer.org' (Lehner et al., 2014).

Data and pedigree information were obtained from the Swedish Dairy Association (Stockholm, Sweden), and the Swedish organic certification organization (KRAV; Uppsala, Sweden) contributed information about dairy farms with organic plant production. ... The initial data set contained records from 402 organic herds (all herds with available data) and 5,335 .... conventional herds (herds with an even last number in the herd identity) (Ahlman et al., 2011).

Explanation

There are many ways eligible study units can be selected, and when multiple organizational levels are used, the

selection approach might differ based on the level. For example, random selection might be used at one level and convenience sampling at another. Clear and transparent descriptions of the selection approach for eligible study units enable identification of the population to which the study results can be inferred and any potential selection biases. When non-probability sampling (e.g. convenience, haphazard or snowball methods) is used, indicate this explicitly and provide a rationale for its use.

6 (c). Describe the method of follow-up

Example 1

After surgery, the owners of the dogs were instructed to monitor for any signs of new mammary tumors and notify the principal investigator (PI) if any signs of recurrence or new tumors were noted. In addition, they were contacted by the PI (VK) every 6 months through phone to ensure this information. ... Dogs with reported/ suspected new tumors were requested to return for clinical examination and confirmation (Kristiansen et al., 2013).

Example 2

Table 1 Possible outcomes of horses on cohort

Possible outcome	Action
No further colic during study	Censored
Colic resolves without medication	Horse returns to population at risk 48 h after colic episode
Colic requires medical attention – clinical records obtained	Horse returns to population at risk 48 h after colic episode
Colic requires surgery	Surgical diagnosis and end of contribution to time at risk
Death from other causes	Censored
Drop out of cohort	Censored/loss to follow-up

(Scantlebury et al., 2011)

Explanation

The potential for loss to follow-up differs between studies; therefore, follow-up monitoring approaches might differ between studies. For example, companion animal populations that rely on client return visits are prone to loss to follow-up, analogous to the human population studies discussed in STROBE. The authors of these studies often make several attempts to contact animal owners to determine their pet's outcome. In other animal populations, data might be collected from computerized systems, such as herd inventory at the start and end of the study, where relevant records (e.g. the reasons for losses) might or might not be available. Reporting the approach used by the authors to minimize loss to follow-up will allow users to assess the potential for bias related to this loss.

6 (d). For matched studies, describe matching criteria and number of matched individuals per subject (e.g. number of controls per case)

Example 1

Two to 4 control farms matched to each case farm on the basis of type of farm (dairy or beef) and location (inside or outside the TB core area) were included in the study. (Kaneene et al., 2002)

Example 2

Each time a herd was recorded as a "case," a randomly selected at-risk herd was identified as a "control". Each control herd was selected with probability proportional to their time at risk (incidence density sampling) during the study period... (Olea-Popelka et al., 2006).

Explanation

Matching is more common in case–control studies, but occasionally, investigators use matching in cohort studies. Matching in cohort studies makes groups directly comparable for potential confounders (Box 5: Confounding) and presents fewer intricacies than with case–control studies. For example, it is not necessary to take the matching into account for the estimation of the relative risk. Because matching in cohort studies might increase statistical precision, investigators might allow for the matching in their analyses and thus obtain narrower confidence intervals.

In case–control studies, matching is done to increase a study's efficiency by ensuring similarity in the distribution of variables between cases and controls, in particular the distribution of potential confounding variables (Costanza, 1995; Sturmer and Brenner, 2002). Example 1 illustrates this type of matching description by matching on farm type and location. Because matching can be done in various ways, with one or more controls per case, the rationale for the choice of matching variables and the details of the method used should be described. Commonly used forms of matching are frequency matching (also called group matching) and individual matching. In frequency matching, investigators choose controls so that the distribution of matching variables becomes identical or similar to that of cases. Individual matching involves matching one or several controls to each case. Matching is not always appropriate in case–control studies, but if used, it needs to be taken into account in the analysis (see Box 2: Matching in case–control studies).

While matching is generally considered to be based on potentially confounding population characteristics, in some case–control studies, the term matching is also used to describe a means of controlling selection from the risk set

Box 5

Confounding

Confounding literally means the confusion of effects. A study might seem to show either an association or no association between an exposure and the risk of a disease. In reality, the seeming association or lack of association is due to another factor that determines the occurrence of the disease but that is also associated with the exposure. The other factor is called the confounding factor or confounder. Confounding thus gives a wrong assessment of the potential 'causal' association of an exposure. For example, an apparent positive association between dogs attending obedience classes and dog bites could occur if specific, large-breed dogs that are prone to biting were more likely to attend the observed obedience classes. In this instance, breed would confound the relationship between obedience class attendance and biting.

Investigators should think beforehand about potential confounding factors, a process that could be enhanced by constructing a causal diagram (see item 7c). An *a priori* consideration of potential confounding variables will inform the study design and allow proper data collection by identifying the confounders for which detailed information should be sought. Restriction, matching or analytical adjustment might also control confounding. In the example above, the study might be restricted to specific breeds. Matching on breed might also be possible, although not necessarily desirable (see Box 2: Matching in case–control studies). There are a number of analytic approaches for identifying confounding variables, which can be broadly grouped into knowledge-based and statistical (Evans et al., 2012; Sauer et al., 2013).

Many of the approaches for controlling confounding assume that the investigator has one or more exposures of interest identified *a priori*. In veterinary literature, observational studies commonly identify risk factors for an outcome from an array of possible independent variables with no *a priori* identification of an exposure of interest or causal diagram.

Regardless of the approach used, when variables are selected for model inclusion, the interpretation of each association needs to be evaluated *post hoc* to evaluate whether all important confounders for that association were included. As part of the *post hoc* assessment, authors should consider whether the variables were confounders or variables with other relationships, such as collider or intervening variables. These other variables can also introduce bias into the association between a different independent variable and the outcome and be detected by algorithm-based approaches (Greenland and Brumback, 2002; Hernan et al., 2002; Schisterman et al., 2009; Hay et al., 2014).

Taking confounders into account is crucial in observational studies, but readers should not assume that analyses adjusted for confounders establish the 'causal part' of an association. Results might still be distorted by residual confounding (the confounding that remains after unsuccessful attempts to control for it), (Olsen and Basso, 1999) random sampling error, selection bias and information bias (see Box 1: Bias in observational studies).

based on the case occurrence timing such as in an incidence density sampling design. Example 2 provides a description of a time-matched selection-control approach.

7 (a). Clearly define all outcomes, exposures, predictors, potential confounders and effect modifiers. If applicable, give diagnostic criteria

Example 1

. . .the explanatory variable of interest was IBK status. Other explanatory variables included in each model as potential effect modifiers or confounders of the association between IBK and weight at ultrasonographic evaluation were birth weight, season, sex of calves after weaning (bull, heifer, or steer), ADG (weaning to yearling weight), preweaning management group, postweaning management group, year of calving, season of calving, the interaction between year and season, and age at ultrasonographic evaluation (Funk et al., 2014).

Example 2

Refer to Section 6(c) for a good description of the outcome event(s) in a cohort study.

Example 3

Body condition was scored from 1 (emaciated) to 5 (obese) using standard methods described by DAFF []. Faecal consistency was scored as described by Alberta Dairy Management [] from 1, representing a liquid consistency, to 4, representing a dry sample. Hide cleanliness was scored following the guidelines of the Food Standards Agency [], where 1 = clean and dry, and 5 = filthy and wet (Williams et al., 2015).

Explanation

Authors should define all variables considered for and included in the analysis, including outcomes, exposures,

Box 2

Matching in case–control studies

In any case–control study, choices need to be made on whether to use matching of controls to cases, and if so, what variables to match on, the precise method of matching to use, and the appropriate method of statistical analysis. Although confounding can be adjusted for in the analysis there could be a major loss in statistical efficiency. The use of matching in case–control studies and its interpretation are fraught with difficulties, especially if matching is attempted on several risk factors, some of which might be linked to the exposure of prime interest (Rothman and Greenland, 1998; Szklo and Nieto, 2000). For example, in a case–control study of a Salmonella outbreak, investigators could match based on factors, such as gender, that are related to the consumption of various food products. However, this control group would no longer represent food consumption choices in the general population, and has several implications. A crude analysis of the data will produce odds ratios that are usually biased towards unity if the matching factor is associated with the exposure. The solution is to perform a matched or stratified analysis (see item 12d). In addition, because the matched control group ceases to be representative for the population at large, the exposure distribution among the controls can no longer be used to estimate the population attributable fraction (see Box 6: Measures of Association and measures of impact) (Cole and MacMahon, 1971). Also, the effect of the matching factor can no longer be studied. If matching is done on multiple factors, the search for well-matched controls can be cumbersome and a non-matched control group might be preferable.

Overmatching is another problem, which might reduce the efficiency of matched case–control studies and, in some situations, introduce bias.

Information is lost and the power of the study is reduced if the matching variable is closely associated with the exposure. Then many individuals in the same matched sets will tend to have identical or similar levels of exposures and therefore not contribute relevant information.

The complexities involved with matching have caused some methodologists to advise against routine matching in case–control studies. Instead, they recommend judicious consideration of each potential matching factor, recognizing that it could potentially be measured and used as an adjustment variable. As a result, studies are reducing the number of matching factors employed, and increasing the use of frequency matching, which avoids some of the problems discussed above. In addition, case–control studies are increasingly abandoning potential confounder matching (Gefeller et al., 1998). Currently, matching remains advisable, or even necessary, when confounder distributions differ radically between the unmatched comparison groups (e.g. age) (Costanza, 1995; Sturmer and Brenner, 2002).

predictors, potential confounders and potential effect modifiers. Disease outcomes require adequately detailed description of the diagnostic criteria. This applies to criteria for cases in a case–control study, disease events during follow-up in a cohort study and prevalent disease in a cross-sectional study.

We advise that authors should declare all 'candidate variables' considered for statistical analysis, rather than selectively reporting only those included in the final models (see also item 16a) (Anderson et al., 2005; Wieland and Dickersin, 2005). Authors should report whether exposures are consistent or change over the study period. For studies involving follow-up, authors should describe how study subjects were uniquely identified, allowing research personnel to correctly record observations at follow-up visits.

7 (b). Describe the level of organization at which each variable was measured

Example

Fixed explanatory variables considered for inclusion in the PA-MNT model were assessment day (d –4, +1, +3, +6, +8, and +10), eye-level IBK-associated corneal ulceration status (present or absent), calf-level IBK-associated corneal ulceration status (present or absent), and landmark (7 levels) (Dewell et al., 2014).

Explanation

Animal populations commonly have multiple organizational levels, so authors should clarify the organizational level at which each variable was measured. For more information, see Box 4: Organization structures in animal populations.

7 (c). For hypothesis-driven studies, the putative causal structure among variables should be described (a diagram is strongly encouraged)

Example

Causal diagrams were constructed to describe postulated links between measured exposure variables and between exposure variables and occurrence of BRD in the first 50 days at risk. As this resulted in a very complex diagram, a

simplified version (only including variables relevant to the assessment of the risk factors included in the analyses reported in this paper) is shown in Fig. 1. Figure 2 shows the causal diagram used to inform the analyses restricted to the three feedlots that routinely used pre-induction assembly. Additional variables included as potential confounders in either of these diagrams were cohort fill duration (all animals added to their cohort within a single day or over a longer period), total number of animals on feed in the animal's feedlot (average for the animal's induction month), number of animals in the animal's cohort, induction weight, breed and season in which the animal was inducted. . . .The DAGitty® software [] was used to identify minimal sufficient adjustment sets to assess total and direct effects of the exposure variable of interest on the occurrence of BRD.

Figures extracted from publication(Hay et al., 2014)

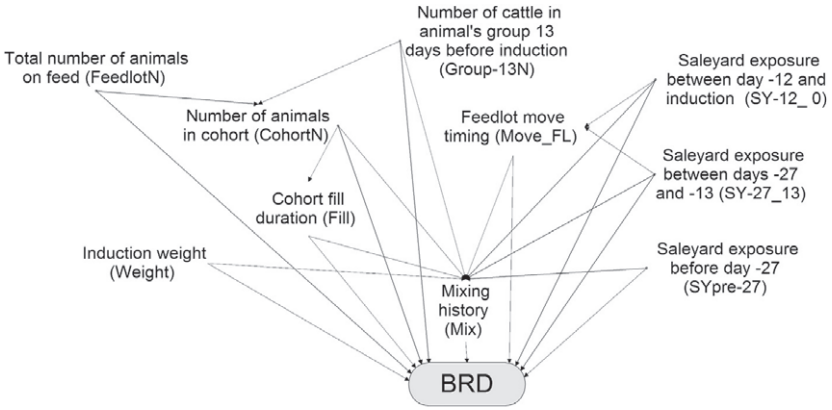

Fig. 1 Causal diagram showing postulated causal paths linking variables related to mixing history, group size, exposure to saleyards and timing of the move to the feedlot to occurrence of bovine respiratory disease (BRD) in the first 50 days on feed.

K.E. Hay , T.S. Barnes , J.M. Morton , A.C.A. Clements , T.J. Mahony

Risk factors for bovine respiratory disease in Australian feedlot cattle: Use of a causal diagram-informed approach to estimate effects of animal mixing and movements before feedlot entry

Preventive Veterinary Medicine, Volume 117, Issue 1, 2014, 160 - 169

http://dx.doi.org/10.1016/j.prevetmed.2014.07.001

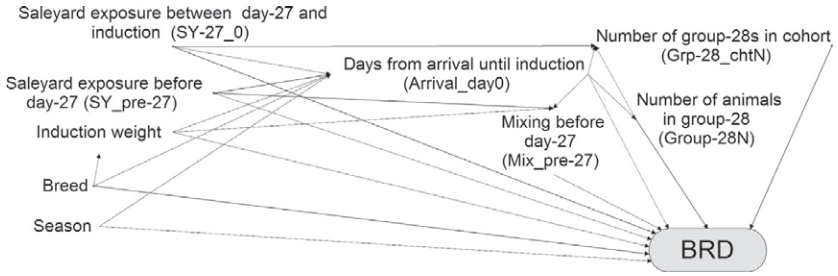

Fig. 2 Causal diagram showing postulated causal paths linking variables related to the interval between arrival and induction, group size and number of groups combined to occurrence of bovine respiratory disease (BRD) in the first 50 days on feed in three ...

K.E. Hay , T.S. Barnes , J.M. Morton , A.C.A. Clements , T.J. Mahony

Risk factors for bovine respiratory disease in Australian feedlot cattle: Use of a causal diagram-informed approach to estimate effects of animal mixing and movements before feedlot entry

Preventive Veterinary Medicine, Volume 117, Issue 1, 2014, 160 - 169

http://dx.doi.org/10.1016/j.prevetmed.2014.07.001

Explanation

For hypothesis-driven studies, it is extremely useful to the end-user if the *a priori* hypothesis and the variable relationships envisioned by the authors are clear and understandable. There are various means available for articulating causal assumptions (Greenland and Brumback, 2002), including directed acyclic graphs (DAGs) (Greenland et al., 1999). Including a causal assumption diagram is strongly recommended. Understanding the underlying causality being explored is important when identifying potential confounding variables and interpreting the results of multivariable analyses. If variables are controlled unnecessarily in a regression model, the power is reduced, and the association between the exposure of interest and the outcome might be biased (Greenland and Brumback, 2002; Hernan et al., 2002; Schisterman et al., 2009).

8 (a). For each variable of interest, give sources of data and details of methods of assessment (measurement). If applicable, describe comparability of assessment methods among groups and over time

Example

Each tumour was examined independently by two specialist veterinary pathologists and, to be included, had to have a minimum of 7 (out of a possible 10) features identified as part of the histopathology study. The 10 features included the presence of: aggregates of lymphocytes, infiltrative margins, intralesional necrosis, perilesional scarring,/inflammation, adjuvant-like material in macrophages, medium-high mitotic rate, giant cells and types of cellular differentiation []. To be included in the estimate of incidence the FISS ("Feline Injection Site Sarcomas" added by authors) had to be diagnosed at the practices for which denominator information was available (Dean et al., 2013).

Explanation

The way in which exposures, confounders and outcomes were measured affects the reliability and validity of a study. Measurement error and misclassification of exposures or outcomes can make it more difficult to detect cause–effect relationships, or might produce spurious relationships. Error in measurement of potential confounders can increase the risk of residual confounding (Becher, 1992; Brenner and Blettner, 1997). It is helpful, therefore, if authors report the findings of any studies of the validity or reliability of assessments or measurements, including details of the reference standard that was used. Rather than simply citing validation studies, we advise that authors give the estimated validity or reliability, which can then be used for measurement error adjustment or sensitivity analyses (see items 12 and 17).

In addition, it is important to know whether groups being compared differed with respect to the way in which the data were collected. For instance, if an interviewer first questions all the cases and then the controls, or vice versa, bias is possible because of the learning curve; solutions such as randomizing the order of interviewing might avoid this problem. Information bias might also arise if the compared groups are not given the same diagnostic

8 (b). If a questionnaire was used to collect data, describe its development, validation and administration

Example

Questionnaire designs were the collective effort of five veterinarians (including four epidemiologists) and a biostatistician. Included in the design group was the Veterinary Officer for Poultry Diseases, who had an in-depth knowledge of each farm as a result of working with the producers to eradicate *Salmonella* from poultry. There were several questionnaires, the main one designed to record independent variables acting at the various levels of broiler production such as at the flock, house and farm levels. During the interval between flocks in each broiler house, a field technician employed by the Veterinary Officer for Poultry Diseases visited each farm to record responses from face-to-face interviews with the person most closely associated with the hands-on management of the broiler flocks and houses, and to record observations of cleaning and disinfection procedures between flocks. The design team reviewed all questions and the method of recording with the field technician to ensure clear understanding. The Veterinary Officer for Poultry Diseases accompanied the field technician on all farm visits and questionnaire recording for the first full month of sampling. During the course of the study, two university-educated field technicians were employed. The first technician was employed for two years, and trained the second technician for one month prior to leaving the project. Interview times varied from 10 to 15 minutes per questionnaire, depending on whether the producer needed to verify records. To ensure consistency in responses, data collected at the previous visit were reviewed with the producer. All questions pertaining to our analysis were closed (Guerin et al., 2007).

Explanation

For STROBE-VET, we needed to draw attention to the descriptions of questionnaire development and administration, because questionnaires are a common data source for veterinary observational studies. Occasionally, authors provide information documenting their questionnaire validation methods, sometimes as a separate publication (Ramon et al., 2008; Dufour et al., 2010). If previous validation information is not available, then

the authors should describe their approach for developing and testing the questionnaire in the manuscript. Like any diagnostic test, unless validated, the diagnostic characteristics of the questionnaire and its ability to accurately measure the variables are unclear. The questionnaire(s) should also be included as supplementary data, or in an open access, permanent site preferably with a Digital Object Identifier (DOI).

8 (c). Describe whether or not individuals involved in data collection were blinded, when applicable

Example

This was an observational study of 292 uniquely identified Bovelder cows born in either 2002 or 2003 (2002 and 2003 cohorts) that were followed from just prior to their first breeding season until they had weaned up to five calves. . . . Farm management and staff were blinded to RTS (reproductive tract scoring) data throughout the study (Holm et al., 2015).

Explanation

While blinding is commonly associated with randomized controlled trials, in observational studies, there is potential for information bias in measurement of exposure arising from knowledge of the outcome of interest (case–control studies) or information bias in measurement of the outcome arising for knowledge of the exposure of interest (cohort studies) (O'Connor et al., 2010; Sargeant et al., 2010). For example, if researchers conduct a case–control study determining factors associated with a tick-borne disease such as Lyme disease (the outcome of interest) and an owner is interviewed about indoor or outdoor exercise (the exposure of interest), the owners of case animals might recall outdoor exercise more easily, because they are familiar with the disease and its causes. This prior knowledge is a potential source of bias. Thus, information about blinding is critical for the reader to assess the impact of bias on the study result. Similar to clinical trials, the use of the terms single- and double blinding should be avoided. Rather, the author should specify the task, caregiver or outcome assessor who is blinded (Giuffrida et al., 2012).

8 (d). Describe any efforts to assess the accuracy of the data (including methods used for 'data cleaning' in primary research, or methods used for validating secondary data)

Example

Selections of dogs from the entire hospital records were made using Oracle programming languages . . . [ ]. First, an in-house hospital code for laboratory-confirmed diagnosis of urolithiasis was used to isolate all the eligible dogs within the boundaries of the study population. . . . Afterwards, urolith laboratory results or medical notes of the identified dogs were manually reviewed to isolate those whose urolith composed of at least 70% monohydrate or dehydrate forms of CaOx (case dogs). Urolith composition was determined at one of two commercial laboratories . . . by means of optical crystallography or infrared spectroscopy as described elsewhere [ ] (Okafor et al., 2014).

Explanation

Reporting the measurement approach is frequently insufficient to ensure validity; therefore, when efforts are made to ensure the data are valid (e.g. the case validity in the example above), these methods should be documented. This documentation enables the end-user to identify potential information bias. In the example above, there could have been concern that the electronic medical records were inaccurate; therefore, the authors validated the electronic medical records by examining the physical medical records, giving the end-user greater confidence in the variable measured. In addition, when data are used for multiple different studies, the data could have been collected for a different purpose initially than that described in the later study. In this case, the original purpose should be described. A description of data validation approaches has recently been published (Emanuelson and Egenvall, 2014).

Bias: describe any efforts to address potential sources of bias due to confounding, selection or information bias

Example

The responses were collected through face-to-face interviews conducted by four experienced interviewers (two teams each comprising two interviewers) between October 2011 and March 2012. As there are different dialects in the Philippines, the questionnaire was written in English and translated to the appropriate dialect at the interview. To reduce information bias the questionnaire was pretested on experts in the Philippines pig production systems comprising regional and provincial veterinary officers and animal health advisors. All questions in the questionnaire were clarified with all interviewers before the study date. The interviewers were instructed to ask questions exactly as stated in the questionnaire and provide only non-directive guidance. To minimize inter-observer variability in conducting the interview, all observers and PVO [Provincial Veterinary Office] personnel met after the questionnaire was piloted on the six farms to agree on a common interpretation of the findings. If there was disagreement, the interpretation of the PVO was chosen. To minimize information (misclassification) and selection biases, the

interviewers were asked to verify the trader's identity, dates when the pigs were sold and number of pigs sold for slaughter before an interview was conducted. . . . The validity of the collected questionnaire data was confirmed during follow-up visits to six farms (three in each province) by the first author, the interviewers and Provincial Veterinary Officers personnel. To reduce misclassification bias that could arise from coding errors, the interviewers and the first author checked and corrected impossible coding of categorical variables (n = 80) and unreliable outlier values for continuous variables (n = 3) (Alawneh et al., 2014).

Explanation

Bias causes study results to differ systematically from the truth. It is important for a reader to know what measures were taken during the conduct of a study to reduce the potential of bias. Ideally, investigators carefully consider potential sources of bias when they plan their study. At the stage of reporting, we recommend that authors always assess the likelihood of relevant biases. Specifically, the direction and magnitude of bias should be discussed and, if possible, estimated. When investigators have set up quality control programmes for data collection to counter a possible 'drift' in measurements of variables in longitudinal studies, or to keep variability at a minimum when multiple observers are used, these should be described. In veterinary medicine, euthanasia or animal culling is a unique potential form of attrition bias, and authors should describe any methods used to account for this bias. Recently, an overview of approaches for addressing bias, including quantitative bias analysis and the use of bias parameters in data analysis, with accompanying veterinary examples was published (Dohoo, 2014).

A discussion about selection bias, information bias and confounding as well as their impact on observational studies is provided in Box 1: Bias in observational studies and Box 5: Confounding .

Box 1

Bias in observational studies

Bias is a systematic deviation of a study's results from a true value. Typically, it is introduced during the design or implementation of a study and its effects cannot be eliminated later or correct analytically. Bias and confounding are not synonymous. Bias arises from flawed information or subject selection so that a wrong association is found. Confounding produces relations that are factually correct, but they cannot be interpreted causally because some underlying, unaccounted for factor is associated with both exposure and outcome (see Box 5: Confounding). Bias differs from random or chance error such as a deviation from a true value caused by random fluctuations in the measured data in either direction. Many potential sources of bias have been described and a various terms have been used (Sackett, 1979). We find that it is helpful to separate them into two simple categories: information bias and selection bias.

Information bias occurs when systematic differences in data completeness or accuracy lead to animal misclassification with respect to exposures, outcomes or measurement errors of values recorded on a continuous scale. Detection bias in cohort studies, interviewer bias and recall bias are all forms of information bias. For example, in a case–control study of risk factors for horse falls, poor dressage performers were less likely to report accurate dressage scores than good performers, thereby introducing information bias (Murray et al., 2004).

Selection bias exists when the association between the exposure and outcome among study-eligible participants is different from those participants included at any stage of the study, from entry to the study to inclusion in the analysis. Various types of selection bias include bias introduced when selecting the control group in a case–control study, differential loss to follow-up, incidence–prevalence bias, volunteer bias, healthy worker bias and non-response bias (Hernán et al., 2004). Detection bias also acts as a form of selection bias in case–control studies (Dohoo et al., 2009).

10 (a). Study size: describe how the study size was arrived at for each relevant level of organization

Example

A sample size of 36 cases and 108 controls was calculated to provide a 95% level of confidence for detecting an odds ratio of 3 with 80% statistical power, assuming a 1:3 ratio of case to control farmers and a random notification process such as a 50% probability of reporting observed oyster mortality. Sample size was increased by 15% to account for non-participation rate observed in previous and recent studies conducted in the same population [], leading to a total of 41 cases and 124 controls, out of 165 and 703 eligible oyster farmers, respectively (Lupo et al., 2014).

Explanation

A study should be large enough to obtain a point estimate with a sufficiently narrow confidence interval to meaningfully answer a research question. Large samples are needed to distinguish a small association from no association. Small studies often provide valuable information, but wide confidence intervals might indicate that they contribute less to current knowledge in comparison with studies providing

estimates with narrower confidence intervals. Also, small studies that show 'interesting' or 'statistically significant' associations are published more frequently than small studies that do not have 'significant' findings. While these studies might provide an early signal in the context of discovery, readers should be informed of their potential weaknesses.

The importance of sample size determination in observational studies depends on the context. If an analysis is performed on data that were already available for other purposes, the main question is whether the analysis of the data will produce results with sufficient statistical precision to contribute substantially to the literature. Formal *a priori* calculation of sample size might be useful when planning a new study (Rigby and Vail, 1998; Carlin and Doyle, 2002). Such calculations are associated with more uncertainty than implied by the single number that is generally produced. For example, estimates of the rate of the event of interest or other assumptions central to calculations are commonly imprecise, if not guesswork (Schulz and Grimes, 2005). The precision obtained in the final analysis can often not be determined beforehand because it will be reduced by inclusion of confounding variables in multivariable analyses (Drescher et al., 1990), the degree of precision with which key variables can be measured, and the exclusion or non-selection of some individuals (Devine and Smith, 1998).

Sample size determination can be complicated further by studies with multiple objectives. Studies frequently have multiple objectives, largely to maximize the amount of data that can be collected from a research study. For instance, a cross-sectional study might estimate an outcome frequency and evaluate the association between one or more exposures on that outcome. It should be clear to the reader which objective was used for sample size determination or, if both objectives were considered, how the final sample size was derived.

In animal health, observational studies might not be hypothesis driven. These studies are not conducted to detect a specific effect size magnitude for an *a priori* identified exposure of interest. Instead, a large number of association measures are calculated with varying levels of precision. This type of study is hypothesis generating. This factor should be discussed specifically, and the rationale for the sample size should be provided. Often, studies do not use formal sample size calculations. For example, when a small number of cases are available for a case–control study, the investigators might choose to include all eligible cases. In this case, the reader still needs to understand how the sample size was derived such as selection of all available cases to evaluate the potential for selection bias or identify an underpowered study.

10 (b). Describe how non-independence of measurements was incorporated into sample size considerations, if applicable

Example 1

The expected prevalence of MRSA was estimated to be considerably lower at 1–2% [], with a much lower between cluster T variance estimated at 0.0001, meaning a total of 800 nasal swab samples would be required to estimate prevalence with a precision of 1% and 95% confidence. To allow for an overall compliance proportion of approximately 60%, each practice was asked to recruit the next 20 horses seen on visits (a total of 1300 horses) (Maddox et al., 2012).

Example 2

Researchers adjusted this sample size[16] for clustering of stillbirth risk in a herd by using the formula $n^* = n[1 + (m - 1)\rho]$, where m is the average herd size, ρ is the intra-class correlation coefficient (ICC), and n is the unadjusted sample size necessary to determine the difference between 2 proportions.[15] Expected herd size was approximately 150 cows and the ICC was estimated to be $\rho = 0.09$ (Waldner, 2008).

Explanation

Given the frequency of non-independent study units in animal populations (see Box 4: Organization structures in animal populations), authors should adjust sample size calculations to account for non-independence. Failure to account for non-independence in sample size determinations might result in studies that are underpowered when analysed correctly using methods that account for clustering. The ethics of conducting underpowered studies are less obvious for observational studies, because study units are observed rather than purposefully assigned to a group. However, resources are potentially wasted when studies are underpowered; therefore, adjustment for non-independence in sample size determinations should be conducted for prospectively planned observational studies.

10 (c). If a formal sample size calculation was used, describe the parameters, assumptions, and methods that were used, including a justification for the effect size selected

Example

. . .prior to conducting the analysis, sample size calculations were performed to determine whether it was likely to obtain a data set of sufficient size to detect a difference of 7.5 kg (16.5 lb) in the primary outcome, live weight, in a

population with 33% of calves in the IBK group and 67% in the unaffected group, with a type I error probability of 0.05, a type II error probability of 0.8, and a 1:2 ratio for case and control calves. The rationale for use of these parameters was that results of a prior study suggested that calves with IBK weighed approximately 7.5 kg less at weaning than unaffected calves, and the prevalence of IBK was approximately 33% in the study herd (Funk et al., 2014).

Explanation

Samples sizes should be calculated based on realistic estimates. While statistical power can be determined using the effect estimate precision and low power affects precision not bias, providing the rationale and assumptions used in the calculations allows the reader to infer the impact of those assumptions on the sample size. For example, what constitutes a meaningful difference might vary between different regions, and the assumed level of non-independence can vary between populations.

11. Quantitative variables: explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why

Examples

Age was grouped on a biological basis into less than 2.5 years, between 2.5 and 4.5 years and more than 4.5 years. This categorisation was decided upon as 2.5 and 4.5 years approximately coincide with ages at first and second parturition in llamas (Rooney et al., 2014).

Explanation

Investigators make choices regarding how to collect and analyse quantitative data about exposures, effect modifiers and confounders. Grouping choices might have important consequences for later analyses (Altman et al., 1994; Royston et al., 2006). We advise that authors explain why and how they grouped quantitative data, including the number of categories, the cut points and category mean or median values (as appropriate). Whenever data are reported in tabular form, the counts of cases, non-cases or controls, animals at risk, animal-time at risk, etc. should be given for each category. Tables should not consist solely of effect-measure estimates or results of model fitting. Authors should state whether categories were selected *a priori* or based on the collected data.

>Investigators might model an exposure as continuous to retain all the information. In making this choice, one needs to consider the nature of the relationship of the exposure to the outcome. Investigators should report how departures from linearity were investigated (e.g. using log transformation, quadratic terms or spline functions). Several methods exist for fitting a nonlinear relation between the exposure and outcome (Greenland, 1995; Royston et al., 1999, 2006). Also, it might be informative to present both continuous and grouped analyses for a quantitative exposure of prime interest.

12. Statistical methods

12 (a). Describe all statistical methods for each objective, at a level of detail sufficient for a knowledgeable reader to replicate the methods. Include a description of the approaches to variable selection, control of confounding and methods used to control for non-independence of observations

Example 1

Collinearity between the variables was investigated by χ^2 analysis. The risk factors initially offered to the model were excluded from the model with a conditional backward elimination procedure; the possible interaction terms were then investigated with a forward conditional selection procedure. A factor was entered in the model at $p \leq 0.05$ and removed at $p \geq 0.10$. The likelihood ratio test was used to assess the overall significance of the model (two-tailed significance level $p \leq 0.05$). Confounding was monitored by evaluating the change in the coefficient of a factor after removing another factor; if the change exceeds 25% of the coefficient value, the removed factor is considered a potential confounder. The significance of each term in the model was tested by Wald's χ^2. In the final model, biologically plausible interaction between factors was investigated by significance. Estimated OR and 95% Wald's confidence interval (CI) were obtained as measures of predictor effect (Rizzo et al., 2016).

Example 2

To account for the hierarchical structure of the data, a cross-classification of feedlot-years (11 feedlots in 2000, 13 in 2001–2002. . .) was included as a random intercept to model the overdispersion arising from the lack of independence of cohorts nested within feedlots, and of feedlots nested within arrival years. In addition, arrival month . . . was modeled as a random intercept using a first-order autoregressive covariance structure to account for the repeated measures of cohorts, within feedlot-years, over months with decay in correlation with increasing distance between observations [] Lastly, arrival week . . . within a month was modeled as a random intercept to control for the correlation of weeks within arrival months (Babcock et al., 2013).

Explanation

Describing statistical methods can be challenging, because the level of detail sufficient for a knowledgeable reader to

replicate the methods is open to interpretation[2] . The author should focus on clearly describing the approach rather than listing statistical tests. Inclusion of a diagram or flow chart to explain a complex analytical process might be helpful. One applicable resource for reporting statistical methods are the SAMPL guidelines (Lang and Altman, 2013). Based on the SAMPL guidelines, the description of the analysis approach can be split into three components: (i) the preliminary analysis, (ii) the primary analysis and (iii) any supplementary analysis. Authors are encouraged to make the data and their software coding available as supplementary material or in data depositories.

In general, there is no one correct statistical analysis but, rather, several possibilities that might address the same question, but make different assumptions. Regardless, investigators should predetermine analyses at least for the primary study objectives in a study protocol. Often additional analyses are needed, either instead of, or as well as, those originally envisaged, and these might sometimes be motivated by the data. Authors should tell readers whether particular analyses were suggested by data inspection. Even though the distinction between pre-specified and exploratory analyses might sometimes be blurred, authors should clarify reasons for particular analyses.

Authors should explain all potential confounders considered, and the criteria for excluding or including variables in statistical models. Decisions about excluding or including variables should be guided by knowledge, or explicit assumptions, on causal relations. Inappropriate decisions might introduce bias, for example by including variables that are in the causal pathway between exposure and disease (unless the aim is to assess how much of the effect is carried by the intermediary variable). If the decision to include a variable in the model was based on the change in the estimate, it is important to report what change was considered sufficiently important to justify its inclusion. If an algorithm such as 'backward elimination' or 'forward inclusion' was used, report the process (including whether a manual or automated process was used) and the significance level and test or other basis (information criteria) for selecting inclusion or exclusion of variable(s) from the model.

12 (b). Describe the rationale for examining subgroups and interactions and the methods used

Example
Biologically important two-way interactions of the explanatory variables in the final model were examined and retained if significant ($p < 0.05$) (Schemann et al., 2011).

Explanation
Subgroup analyses and interactions can be planned or conducted after reviewing the data. Authors should report if the subgroup analysis was pre-planned or informed by data examination. This information allows the end-user to identify the presented associations in the context of hypothesis testing or hypothesis generating.

12 (c). Explain how missing data were addressed

Examples
In model 1, only subjects with complete information on variables in the final model were included. Model 2 was a Bayesian full-likelihood analysis where missing data were taken into account and became a multidimensional additional parameter [] (Rossow et al., 2014).

Explanation
Missing data are common in observational research. Questionnaires are not always filled in completely, owners might not bring their animal to all follow-up visits and routine data sources and clinical databases are often incomplete. For analyses that account for missing data, authors should describe the nature of the analysis (e.g. multiple imputation) and the assumptions that were made (See Box 7: Missing data: problems and possible solutions).

In cases where euthanized or culled animals are designated as missing data or observations lost to follow-up in the analysis, the authors should clearly identify this criterion as a cause of missing data. Describing this aspect of the analysis is important, because the circumstances surrounding culling or euthanasia are likely not to be random, which violates the assumption that the missing data are random. The approach for reporting missing observations or loss to follow-up is discussed in item 13.

12 (d). If applicable, describe the analytical approach to loss to follow-up, matching, complex sampling and multiplicity of analyses

Example
Conditional logistic regression [] was used to assess differences in mean production (3.5% FCM, fat, protein, LSCC), JD test status as adults, removal from herd during the observation period, and JD test status of dam between cases and controls (Pillars et al., 2011).

Explanation
For cohort studies, authors should report whether they conducted analyses to determine whether loss to follow-up was differentially associated with other factors. Another consideration is the approach for handling

[2]http://cdn.elsevier.com/promis_misc/AMEPRE_gfa_mar2015.pdf

Box 7

Missing data: problems and possible solutions

Missing data are common in observational research. In studies conducted in populations with multiple organizational levels, missing data might occur and need to be described at multiple levels. Rubin developed a typology of missing data problems, based on a model for the probability of an observation being missing (Rubin, 1976; Little and Rubin, 2002). Data are described as missing completely at random (MCAR) if the probability that a particular observation is missing does not depend on the value of any observable variable(s). Data are missing at random (MAR) if, given the observed data, the probability that observations are missing is independent of the actual values of the missing data. For example, suppose younger dairy heifers are more prone to missing pregnancy checks, but the probability of missing the check is unrelated to the true pregnancy risk after accounting for age. Then the missing pregnancy measurements would be MAR in models including age. Data are missing not at random (MNAR) if the probability of missing still depends on the missing value even after taking the available data into account. When data are MNAR, valid inferences require explicit assumptions about the mechanisms that led to missing data. In studies with multiple organizational levels, data might be missing at the individual level, group level or both. For example, the probability of loss to follow-up might depend on both group and individual characteristics (Diaz-Ordaz et al., 2014).

Methods to deal with data missing at random (MAR) fall into three broad classes: (Rubin, 1976; Little and Rubin, 2002) likelihood-based approaches (Lipsitz et al., 1999), weighted estimations (Rotnitzky and Robins, 1997) and multiple imputation (Schafer, 1997; Donders et al., 2006; Rubin, 2008). Options for dealing with missing data in veterinary literature have recently been published, along with an assessment to determine the magnitude of bias that might arise from a complete-case analysis (Dohoo, 2015).

failure to observe the outcome, which can vary greatly depending upon the disease frequency measure such as rates or risk (Box 6: Measures of Association and measures of impact). In cohort studies that use rates such as animal-time at risk), the observed time of animals lost to follow-up are included in the analysis, and the outcome is censored. Because the approach used to analyse censored data also varies, it should be described accordingly (see item 7). When performing a survival analysis, an unobserved outcome can be the result of loss to follow-up or completion of the study. The authors should clearly state whether the analysis treats these two forms of censoring differently. When incidence risk (cumulative

Box 6

Measures of association and measures of impact

The terms used to describe metrics in epidemiology are, unfortunately, not consistent. Therefore, care is needed when deciding which concept is being described and if it is appropriate for a particular situation. For example, Dohoo et al. (2009) and the Centers for Disease Control Principles of Epidemiology in Public Health Practice (Anonymous, 2012) use the term Measures of Association to describe measures such as odds ratio, rate ratio, while the on-line text book Active Epi (http://activepi.herokuapp.com/courses/active-epi-course) uses the term Measure of Effect to describe these measures. Regardless, of the broad grouping name used, these measures are usually the ratio of two measures of disease frequency. The relative measures emphasize the strength of an association, and are most useful in aetiological research.

In addition, another set of measures are those are designed to answer the question 'How much of the disease burden in a population could be prevented by eliminating the exposure?' The category used to describe these measures again differs by author. Dohoo et al. (Dohoo et al., 2009) uses the term 'Measures of Effect', the CDC Manual for Epidemiology uses the term 'Measures of Public Health Impact' and Kleinbaum (Kleinbaum, 2009) uses the term 'Measures of Potential Impact'. These calculations cover several concepts (and no unifying terminology) exist, and incorrect approaches to adjust for other factors are sometimes used (Greenland, 1998; Uter and Pfahlberg, 2001). For example, Kleinbaum suggests that the terms risk difference (RD), attributable risk and excess risk are synonyms (Kleinbaum, 2009). Similarly, Klieinbaum suggests that aetiological fraction (EF) in the population can also be called the population attributable risk, the population attributable risk per cent, the population attributable risk proportion and the population attributable risk fraction. The EF is appropriate for cohort studies that estimate risk such as cumulative incidence. If the study measures incidence rate, these terms change accordingly. Another measure is the aetiological fraction among exposed (EFe), which is alternatively called the attributable risk per cent among exposed, the attributable risk fraction among the exposed and the attributable risk proportion among the exposed. This measure focuses on the potential impact of the exposure on the number of exposed cases, rather than the total number of cases of the disease. Regardless of the number used, authors should be aware of the strong assumptions made in this context of using aetiological fractions (Rockhill et al., 1998).

incidence) is the disease frequency measure, authors should explain how they interpreted data about animals that leave the study before the end of the study

In individually matched case–control studies, a crude analysis of the odds ratio ignoring the matching usually leads to an estimation that is biased towards unity (see Box 2: Matching in case–control studies).

When authors use complex, multiple-stage sampling schemes to select the study population, authors should describe how this scheme is incorporated into the data analysis, thereby providing a valid estimate of effect size and precision.

When authors conduct multiple hypothesis tests, then authors should indicate if they did or did not use a method to adjust the definition of a 'statistically significant' *P*-value. The description of the method should clarify whether an adjustment approach for multiple comparisons was employed within a specific hypothesis test.

12 (e). Describe any analyses used to assess the robustness of the analyses (e.g. sensitivity analyses or quantitative bias assessment)

Example

The national database used to sample controls did not enable us to take into account the size of the flocks. Therefore, counties with a large percentage of small flocks (<20 ewes) might have been overrepresented. To assess the influence of geographic selection bias, we conducted a sensitivity analysis by using 2 methods: 1) weighting of controls in the final model with weights being defined for each county as the ratio of the percentage of flocks >20 ewes in the county divided by the percentage of flocks >20 ewes at the national scale, and 2) introduction of sheep production areas as random coefficients in the final model (Fediaevsky et al., 2009).

Explanation

Sensitivity analyses are useful to investigate whether or not the main results are consistent with those obtained with alternative analysis strategies or assumptions (Rothman and Greenland, 1998). Issues that might be examined include the criteria for inclusion in analyses, the definitions of exposures or outcomes (Custer et al., 2006), which confounding variables merit adjustment, the handling of missing data (Wakefield et al., 2000; Dunn et al., 2001), possible selection bias or bias from inaccurate or inconsistent measurement of exposure, disease and other variables, and specific analysis choices, such as the treatment of quantitative variables (see item 11). Sophisticated methods are used increasingly to simultaneously model the influence of several biases or assumptions (Greenland, 2003; Lash and Fink, 2003; Phillips, 2003).

Results

The results section should give a factual account of what was found, from the recruitment of owners/managers and their animals and the description of the study populations to the main results and ancillary analyses. The results should be reported in sufficient detail for secondary use of the data (e.g. for meta-analysis or risk assessment). The results section should be free of interpretations and discursive or overly discussive text reflecting the authors' views and opinions.

13. Participants

13 (a). Report the numbers of owners / managers and animals at each stage of study and at each relevant level of organization –for example, numbers eligible, included in the study, completing follow-up and analysed

Example 1

During the study period, a total of 2457 cats attended the Small Animal Teaching Hospital. Of these, 237 records were identified after the database search, and 174 cases met the eligibility criteria (Trehy et al., 2014).

Example 2

Table 1 Structure of the data from 3027 lactation records from dairy cows on Reunion Island (1993–1996)

Level	Number	Average number per unit at next-higher level	Range
Region (highest level)	5	–	–
Herd	50	10	3–16
Cow	1570	31.4	8–105
Lactation	3027	1.9	1–5

(Dohoo et al., 2001)

Explanation

Detailed information on the process of recruiting study participants is important for several reasons. Those included in a study often differ in relevant ways from the target population to which results are applied. This might result in estimates of prevalence or incidence that do not reflect the experience of the target population and lead to selection bias (see Box 1: Bias in observational studies).

Investigators should give an account of the numbers of owners/managers and animals considered at each stage of recruiting study participants and at each level of organization. The choice of a target population and the detailed criteria for inclusion of participants' data in the analysis should be described. Depending on the type of study, this might include the number of owners/managers and animals found to be eligible, the number included in the study, the number examined, the number followed up and the number included in the analysis. Information on different organizational levels might be required, if sampling

of study participants is carried out at two or more organizational levels (multistage sampling). In case–control studies, we advise that authors describe the flow of participants separately for case and control groups. (Schulz and Grimes, 2002) Controls can sometimes be selected from several sources, including veterinary clinics and community dwellers.

13 (b). Give reasons for non-participation at each stage and at each relevant level of organization

Example 1

We investigated a total of 233 known OJD infected flocks to identify eligible flocks, of which the eligibility of 32 (13.7%) could not be determined because the farmer refused to participate for various reasons (lack of interest (6), old age or health problems (4), inability to muster sheep (2), anger about past surveys (1) and no reason given (19)) (Dhand et al., 2007).

Example 2

Reasons for exclusion were lack of a result for serum cobalamin (35 cats), cobalamin measured using a different method (13 cats), or incomplete clinical records (15 cats). A record of prior cobalamin supplementation (within the three months before presentation) was identified in 18 cases (16 with serum cobalamin greater than the reference interval, 1 with cobalamin within the reference interval and 1 with cobalamin below the reference interval) (Trehy et al., 2014).

Explanation

Although low participation does not necessarily compromise the validity of a study, transparent information on participation and reasons for non-participation are essential. Also, as there are no universally agreed definitions for participation, response or follow-up rates, readers need to understand how authors calculated such proportions (Slattery et al., 1995; Galea and Tracy, 2007). Explaining the reasons why owners/managers or animals no longer participated in a study or why they were excluded from statistical analyses helps readers judge whether the study population was representative of the target population and whether bias was possibly introduced. For example, in a survey of horse owners investigating an equine health outcome, non-participation due to reasons unrelated to a horses' health status (such as the survey not being delivered due to an incorrect address) might affect the estimate precision but is not likely to introduce bias. Conversely, if owners/managers opt out of the survey because their horse is ill or perceived to be in excellent health, the results might underestimate or overestimate the population's prevalence of ill health. If failure to participate or loss to follow-up during the study is related to both an exposure of interest and the outcome, the relationship between the exposure and the outcome might also be biased.

13 (c). Consider use of a flow diagram and/or a diagram of the organizational structure

Example

Figure extracted from publication(Menzies et al., 2012)

Explanation

An informative and well-structured flow diagram can readily and transparently convey information that might otherwise require a lengthy description (Egger et al., 2001). The diagram might usefully include the main results such as the number of events for the primary outcome. The flow chart might need to include information for both owners/managers and animals as well as information at multiple organizational levels, if applicable.

14. Descriptive data on exposures and potential confounders

14 (a). Give characteristics of study participants (e.g. Demographic, clinical, social) and information on exposures and potential confounders by group and level of organization, if applicable

Example

Table 2 Descriptive statistics for variables of interest in calves born in Fall 2005-2008

Year	Pinkeye: n	ADG: kg/d (SD)	URFAT: cm (SD)	UFAT: cm (SD)	UREA: cm^2 (SD)	UPFAT: % (SD)	LIVEWT: kg (SD)
2005	Case: 15	0.92 (0.37)	0.58 (0.38)	0.53 (0.36)	55.13 (17.55)	4.45 (1.02)	344.40 (98.91)
	Neg: 92	1.01 (0.40)	0.66 (0.34)	0.62 (0.38)	60.21 (16.48)	4.72 (1.12)	382.59 (100.75)
2007	Case: 9	0.98 (0.55)	0.60 (0.42)	0.54 (0.43)	59.90 (22.45)	4.27 (0.99)	380.16 (140.17)
	Neg: 123	0.98 (0.49)	0.65 (0.39)	0.59 (0.35)	60.69 (18.23)	4.55 (0.98)	381.99 (116.58)
2008	Case: 3	0.76 (0.39)	0.70 (0.45)	0.62 (0.34)	53.10 (18.26)	4.55 (0.32)	353.95 (120.51)
	Neg: 126	1.02 (0.37)	0.71 (0.29)	0.65 (0.31)	65.77 (17.80)	4.76 (1.15)	406.73 (96.51)
Total	Case: 27	0.92 (0.42)	0.59 (0.39)	0.55 (0.37)	56.50 (18.79)	4.41 (0.94)	357.38 (112.73)
	Neg: 341	1.00 (0.42)	0.67 (0.34)	0.62 (0.34)	62.44 (17.75)	4.63 (1.09)	391.29 (105.63)

(Funk et al., 2014)

Example 2

Table 1 Descriptive statistics (mean, (SD, median, min, max)) for PA-MNT (Pressure Algometry- Mechanical Nociceptive Threshold in kg/f) for calves scarified on d 0 (n=number of eyes)

d (n)	Landmark 1	Landmark 2	Landmark 3	Landmark 4 (Control)
Scarified eyes				
−4 (19)	5.2 (1.9, 4.8, 1.4, 10.1)	2.0 (1.2, 1.8, 0.2, 6.5)	4.0 (2.4, 3.7, 0.5, 10.3)	5.5 (2.2, 5.2, 1.1, 11.1)
1 (19)	3.9 (1.9, 3.7, 0.7, 10.0)	1.7 (1.1, 1.4, 0.1, 5.1)	3.1 (2.1, 2.5, 0.1, 8.3)	3.9 (1.7, 3.9, 0.7, 11.2)
3 (16)	3.1 (1.2, 3.0, 0.9, 6.9)	1.3 (0.8, 1.2, 0.2, 3.3)	2.6 (2.5, 1.6, 0.2, 11.2)	3.2 (1.1, 3.2, 0.4, 6.6)
balance of table omitted. . .				

(Dewell et al., 2014)

Explanation

Readers need descriptions of study participants and their exposures to judge the generalizability of the findings or use the data in secondary analyses. In veterinary studies, this might include descriptive information about the owners/managers, herds, pens and animals. In studies that compare groups, the descriptive characteristics and numbers should be given by group. The 'group' variable would be exposure level or outcome status, depending on the study subject selection method. Inferential measures such as standard errors and confidence intervals should not be used to describe the variability of characteristics, and significance tests and *P*-values should be avoided when describing the baseline characteristics of the study population. In cohort studies, it might be useful to document how an exposure relates to other characteristics and potential confounders. Authors could present this information in a table with columns for participants in two or more exposure categories, which permits the reader to judge the differences in confounders between these categories.

Information about potential confounders, including whether and how they were measured, influences judgments about study validity. We advise authors to summarize continuous variables for each study group by giving the mean and standard deviation, or, when the data have an asymmetrical distribution (as is often the case), the median and percentile range (e.g. 25th and 75th percentiles). Variables made up by small number of ordered categories (such as stages of disease I to IV) should not be presented as continuous variables; it is preferable to give numbers and proportions for each category. The SAMPL guidelines provide recommendations for reporting descriptive statistics for different variable types (Lang and Altman, 2013). We recommended that descriptive information be provided for all variables measured in the study, regardless of whether they are included in the final analyses. To allow the reader to evaluate the statistical power for an individual variable and the probability of a type I error given the total number of variables evaluated, authors should provide information on the number of variables and the distribution of data among each variable's categories. Some journals might be reluctant to publish extensive descriptive tables due to word limits or page constraints. In these cases, we recommend that the descriptive information be provided as supplementary material and the total number of associations tested be provided in the main text. The approach for presenting inferential statistics is discussed in item 15. The decision to combine descriptive statistics (item 13) and inferential statistics (item 15), into one table, as was done in the example provided for this item, depends on author and journal preference.

14 (b). Indicate number of participants with missing data for each variable of interest and at all relevant levels of organization

Example

In total 112 farmers returned useable prospective records on 2143 litters. . . . Cause and timing of piglet death data were returned for 2826 piglets from 1304 litters from 111 farms. . . . With the exception of unknown sow parity, incomplete piglet mortality records were excluded from the risk factor analysis, this amounted to 1714 piglet records. Data on a cohort of 25,761 piglets from 2143 litters from 112 farms were analysed (KilBride et al., 2012).

Explanation

As missing data might bias or affect generalizability of results, authors should tell readers the amounts of missing data for exposures, potential confounders and other important characteristics of study subjects (see item 12c and Box 7: Missing data: problems and possible solutions).

F.D. Menzies et al./The Veterinary Journal 194 (2012) 299–302

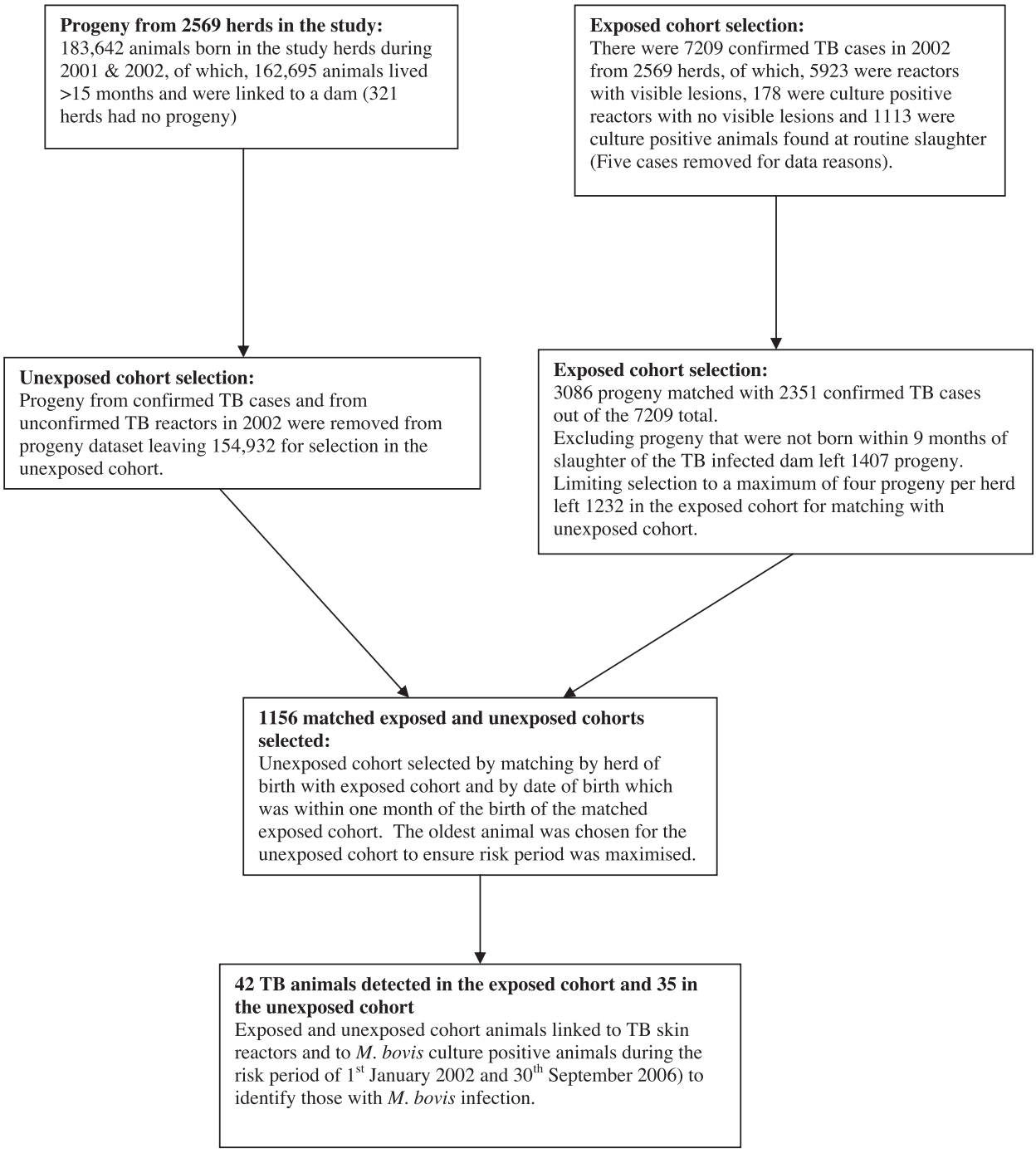

Fig. 1. Tuberculosis (TB) cohort selection data generation tree for the study.

Authors should clearly describe the number of animals missing due to elective euthanasia or culling. Authors also should report numbers at each level or organization. A study with a small number of missing observations from each herd might have different implications than a study where all of the missing data are from one herd. We advise authors to use their tables and figures to enumerate amounts of missing data.

14 (c). Summarise follow-up time (e.g. Average and total amount), if appropriate to the study design

Example

A total of 548 calves were recruited and followed up to 51 weeks or until they died, contributing a total of 25,104 calf weeks (481.1 calf years) of life to the study. Five animals were lost to follow up due to non-compliance to study protocol or were stolen from the study farms. A total of 88 calves died before reaching 51 weeks of age, giving a crude mortality rate of 16.1 (13.0–19.2; 95% CI) per 100 calves in their first year of life. Of the 88 animals that died, 33 deaths were attributed to East Coast fever, 10 to haemonchosis, and 6 to heartwater (Thumbi et al., 2014).

Explanation

Readers need to know the duration and extent of follow-up for the available outcome data. Authors can present a summary measure of the follow-up such as the mean follow-up time, median follow-up time or both, as appropriate. The mean allows a reader to estimate the total number of animal-years by multiplying it with the number of study subjects. Authors also might present minimum and maximum times or percentiles of the distribution to show readers the spread of follow-up times. They might report total animal-years of follow-up or some indication of the proportion of potential data that were captured (Clark et al., 2002). All such information might be presented separately for animals in two or more exposure categories.

15. Outcomes

Example 1

In the reduced dataset of 200 cats, 126 (63%) cats were classed as overweight and 74 (37%) as normal weight. The BCS distribution for the study population without the cats that were excluded on medical grounds (n = 206) as shown in . . ..

Table 3 Variables assessed as potential risk factors for being obese or overweight (O/wgt) grouped according to model of 200 cats included in the study (New Zealand 2007).

Model	Variable	Category levels	Cats (n)	O/wgt (%)	p-value
Cat characteristics	Age (years)[a]	≤2	29	41	0.002
		3–7	87	60	
		8–12	58	69	
		≥13	24	88	
	Desexed?[a]	Entire	9	11	0.002
		Desexed	190	65	
	Gender	Male	109	66	0.463
		Female	92	60	
. . . balance of table omitted					

(Cave et al., 2012)

Example 2

Table 4 Multivariable model of risk factors for the occurrence of Campylobacter at first-depopulation in 354 conventional broiler flocks in Northern Ireland, June 2001 to May 2002

Variable	No. of flocks	Positive (%)	Odds ratio	95% CI	p-Value (Wald's)	p-Value of factor
Age at samplinga						
Per day increase	354	42.9	1.16	1.05–1.28	0.005	0.004
Number of houses on site						
One	125	31.2	1	–	–	0.018
Two	88	39.8	1.39	0.60–3.21	0.447	–
Three or more	141	55.3	2.86	1.32–6.22	0.008	–
. . . balance of table omitted						

(McDowell et al., 2008)

Example 3

The FEC results were reported as ep5 g [eggs per 5 grams of feces]. The range of egg counts was from 0 to 419 ep5 g. Table III summarizes the range, mean, and median counts for the 4 regions.

Table III Summary statistics for 1947 fecal egg counts of Trichostrongle-type eggs per 5 g of feces (ep5 g) from samples collected from 38 Canadian dairy herds.

	Range	Mean	Median	SD	Variance	N
PEI	0–419	12.8	2	37.8	1428.5	1016
Quebec	0–241	7.8	1	23.6	558.7	610
Ontario	0–48	2.2	0	6.1	37.2	163
Saskatchewan	0–189	5.6	0	25.5	652.7	157
Overall	0–419	9.8	1	29.0	998.7	1946

SD – standard deviation

(Nødtvedt et al., 2002)

Explanation

Before addressing the possible association between exposures (risk factors) and outcomes, authors should report relevant descriptive data. It might be possible and meaningful to present unconditional measures of association in the same table that presents the descriptive data.

15 (a). Report outcomes as appropriate for the study design and summarize at all relevant levels of organization

Item 15 differs from item 14, in that 15 explicitly relates to the outcome (event) information. In cross-sectional and risk-based cohort studies, authors should report the number of events for each outcome of interest. For example, in Example 1, this information is provided in the table. Consider presenting this information separately for participants

in different categories of key exposures of interest. Example 1 also includes information relevant to item 14 such as information about the distribution of potential confounders. For rate-based cohort (longitudinal) studies, consider reporting the event rate per animal-year of follow-up. For case–control studies, the focus will be on reporting exposures separately for cases and controls as frequencies or quantitative summaries.

Describing the outcome at all organizational levels requires balancing between complete reporting and 'information overload'. If the outcome analyses are all carried out at the lowest level, the outcome at all higher levels might not be need to be reported. However, the authors should provide the reader with some idea as to how the outcome varies across higher level units. In Example 3, the hierarchy consisted of province ($n = 4$), herd ($n = 38$), cow ($n = 304$) and sample ($n = 1946$). The authors provided some evidence of the variability in faecal egg count across provinces by providing descriptive statistics in a tabular form by province. In some cases, it might be appropriate to report the outcome at different time points (e.g. for a longitudinal study).

In Example 4, the organizational levels were herd ($n = 210$) and cows ($n \approx 5000$). To indicate the range of incidence rates of several diseases across herds, they presented those rates as box-and-whisker plots, which effectively convey the cross-herd variability. When figures are used for presentational clarity, tables of numerical values, which are required for meta-analyses and risk analyses, should be included in the main text or supplemental material.

Multilevel studies are often analysed using random-effects models. In these cases, the authors should present the variance estimates at all levels to provide information about the outcome variability across all organizational levels.

15 (b). For proportions and rates, report the numerator and denominator

It is important to present both numerator and denominator values, so users can calculate unconditional risk ratios (RR) or odds ratios (OR). In Example 1 (cross-sectional study), the number of animals in the exposure variable categories and the per cent with the outcome are reported. These values allow the reader to compute the numerator and denominator values for a RR. In Example 2 (case–control study), the actual numerator and denominator values (listed by key exposure variable categories) and the associated ORs are presented. For rate-based studies, the number of outcome events and amount of animal-time at risk should be presented for key predictor categories.

15 (c). For continuous outcomes, report the number of observations and a measure of variability

For quantitative outcomes, present appropriate summary measures. For (approximately) normally distributed values, the authors should report the mean and standard deviation (SD) or variance. We do not recommend reporting the standard error of the mean, because standard error is an inferential statistic rather than a descriptive one. For non-normally distributed outcomes, report either the mean and SD of a normally distributed transformed outcome or consider reporting the median and interquartile range (or complete range) of the original variable. In Example 3, the mean and SD as well as the median and range have been presented.

16. Main results

16 (a). Give unadjusted estimates and, if applicable, adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders and interactions were adjusted. Report all relevant parameters that were part of the model

Example 1
(Willeberg et al., 2012)

Table 2 Example 1: model estimates of the effects of alternative confounder adjustments based on data from Table 1

# of cattle purchased from positive herds in the previous quarter	Full model OR_f estimates[a]	Simplified model OR_f estimates			Crude OR_f
		1[b]	2[c]	3[d]	
>20	7.4	8.1	9.7	9.9	9.7
10-20	11.0	11.9	14.2	13.8	14.2
1-10	3.8	3.8	5.0	4.8	5.0
. . . balance of table omitted					

For details on the multivariable model, see (Nielsen et al., 2007)
[a]Based on the full model with four significant risk factors as presented by Nielsen et al. (2007).
[b]Model 1: as the full model without the risk factor: 'herd size'.
[c]Model 2: as Model 1 without the risk factor: 'region of the country'.
[d]Model 3: as Model 2 without the risk factor: 'number of *Salmonella*-positive herds in the previous quarter within a 2 km radius', i.e., model with the primary risk factor, controlling for repeated measurements within herds.

Example 2

(Pires et al., 2013)

Table 3 Final multivariable random effects logistic regression models of associations between barn thermal environment parameters, pig-, pen- and cohort-level risk factors and Salmonella shedding in finishing pigs in three sites. Multilevel multivariable logistic models with random intercepts at pig-, pen- and cohort-levels.

Models	Measured at level	Independent variable	Beta	SE	OR	95% CI	*p*-Value
Model 1		Intercept	−2.65	0.58	–	–	–
	Pig	Age[f]	−1.18	0.017	0.7	0.65–0.74	<0.001
	Pen	Cold exposure 12h[i]	0.44	0.2	1.51	1.02–2.25	0.03
	Cohort	Nursery status[g]	2.16	0.52	4.14	2.79–17.15	<0.001
	Farm	Site					
		A vs B	0.93	0.61	2.52	0.76–8.42	
		A vs C	0.38	0.63	1.46	0.42–5.04	
		B vs C	1.3	0.64	3.69	1.06–12.86	

. . . balance of table and footnotes omitted

Cohort (Var^h = 0.77 (0.43), % = 11.48); Pen (Var^h = 1.69 (0.34), % = 25.19); Pig (Var^h = 0.96 (0.21), % = 14.31); Total (Var^h = 6.71)
Cohorts (n = 18); Pens (n = 361); Pigs (n = 899); Individual fecal samples (n = 6751); *Salmonella* prevalence (6.58%)
h(Variance components, standard error and proportion of variance at the cohort-, pen- and pig-level. Individual fecal sample variance: $\pi^2/3 = 3.29$ (latent-variable technique).

Explanation

In many situations, authors might present the results of unadjusted or minimally adjusted analyses and those from fully adjusted analyses. We advise giving the unadjusted analyses together with the main data, for example the number of cases and controls that were exposed or not. This allows the reader to understand the data behind the measures of association (see also item 15). For adjusted analyses, report the number of animals in the analysis, as this number might differ because of missing values in covariates (see also item 12c). Estimates should be given with confidence intervals.

Readers can compare unadjusted measures of association with those adjusted for potential confounders and judge by how much, and in what direction, they changed. Readers might think that 'adjusted' results equal the causal part of the measure of association, but adjusted results are not necessarily free of random sampling error, selection bias, information bias or residual confounding. Thus, great care should be exercised when interpreting adjusted results, as the validity of results often depends crucially on complete knowledge of important confounders, their precise measurement and appropriate specification in the statistical model (see also item 20) (Smith and Phillips, 1990; Christenfeld et al., 2004).

Data non-independence is frequently encountered in animal studies and often addressed by fitting a random-effects model. It is important that these random effects be reported (and interpreted), because they are as important a model component as the fixed effects (see Example 2).

16 (b). Present category boundaries when continuous variables were categorized

Explanation

Categorizing continuous data has several important implications for analysis (Box 3: Grouping / C) and also affects the presentation of results. In tables, outcomes should be given for each exposure category, for example as counts of animals at risk, animal-time at risk, if relevant separately for each group (e.g. cases and controls). Details of the categories used might aid comparison of studies and meta-analysis. If data were grouped using conventional cut points (e.g. below normal, normal reference range, above normal for body temperature), group boundaries such as range of values can be derived easily, except for the highest and lowest categories. If quantile-derived categories are used, the category boundaries cannot be inferred from the data. At a minimum, authors should report the category boundaries; it is helpful also to report the range of the data and the mean or median values within categories. Commonly, category boundaries are presented directly in the table with model results (see Example 1).

Box 3

Grouping / Categorization

There are several reasons why continuous data might be grouped (Altman, 2005). When collecting data, it might be better to use an ordinal variable than to seek an artificially precise continuous measure for an exposure based on recall over several years. Categories might also be helpful for presentation, for example to present all variables in a similar style, or to show a dose–response relationship.

Grouping might also be done to simplify the analysis, for example to avoid an assumption of linearity or when investigating interactions between two continuous variables. However, grouping loses information and might reduce statistical power (Cohen, 1983) especially when dichotomization is used (Zhao and Kolonel, 1992; MacCallum et al., 2002; Royston et al., 2006). If a continuous confounder is grouped, residual confounding might occur, whereby some of the variable's confounding effect remains unadjusted (see Box 5: Confounding) (Cochran, 1968; Becher, 1992). Increasing the number of categories can diminish power loss and residual confounding, and is especially appropriate in large studies. Small studies might use few groups because of limited numbers.

Investigators might choose cut points for groupings based on commonly used values that are relevant for diagnosis or prognosis, for practicality, or on statistical grounds. They might choose equal numbers of individuals in each group using quantiles (Clayton et al., 1993). On the other hand, one might gain more insight into the association with the outcome by choosing more extreme outer groups and having the middle group(s) larger than the outer groups (Cox, 1957). In case–control studies, deriving a distribution from the control group is preferred as it is intended to reflect the source population. Readers should be informed if cut points were selected *post hoc*. In particular, if the cut points were chosen to minimize a *P* value, the true strength of an association will be exaggerated (Altman et al., 1994).

When analysing grouped variables, it is important to recognize their underlying continuous nature. For instance, a possible trend in risk across ordered groups can be investigated. A common approach is to model the rank of the groups as a continuous variable. Such linearity across group scores will approximate an actual linear relation if groups are equally spaced but not otherwise. Il'ysova et al. (Il'yasova et al., 2005) recommends publication of both the categorical and the continuous estimates of effect, with their standard errors, to facilitate meta-analysis, as well as providing intrinsically valuable information on dose–response. One analysis might inform the other and neither is assumption-free. Authors often ignore the ordering and consider the estimates (and *P*-values) separately for each category compared with the reference category. This might be useful for description, but might fail to detect a real trend in risk across groups. Recent method developments, such as fractional polynomials that fit a wide range of nonlinear relationships (Royston and Sauerbrei, 2008), and the availability of software to implement these methods in standard software packages reduces the need to routinely categorize variables.

16 (c). If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period

Example

Table 4 Population attributable fraction (PAF) and 95% confidence interval for selected explanatory variables regarding 295 dog owners in Taiwan (2004)

		95% CI	
	PAF (%)	Lower limit	Upper limit
History of unsuccessful ownership	33	11	50
Spayed after giving birth [a]	22	3	37
Soiling	17	6	26
Barking	13	1	23
Barking and soiling combined	23	3	40
Neighbor complaints [b]	11	1	20

[a]Do you think that a female dog would be healthier if she had one litter before being fixed?

[b]Have there been any neighborhood problems or complaints concerning your dog?

(Weng et al., 2006)

Explanation

In many circumstances, the absolute risk associated with an exposure is of greater interest than the relative risk. Absolute measures such as risk different and measures of impact such as aetiological fraction in the exposed or aetiological fraction in the population might be useful to gauge how much disease can be prevented if the exposure is eliminated (example above). For this reason, such measures are of interest to the end-users. These measures should preferably be presented together with a measure of statistical uncertainty, although CI calculations might be difficult for estimates derived from multivariable models. Authors should be aware of the strong assumptions made in this context, including a causal relationship between a risk factor and disease (also see Box 6: Measures of Association and measures of impact) (Rockhill et al., 1998). Because of the semantic ambiguity and complexities involved, authors should report the method used to calculate such measures, ideally giving

Box 8

Interaction (effect modification): the analysis of joint effects

Interaction exists when the association of an exposure with the risk of disease differs in the presence of another exposure. One problem in evaluating and reporting interactions is that the effect of an exposure can be measured in two ways: as a risk ratio (or rate ratio) or as a risk difference (or rate difference). The use of the ratio leads to a multiplicative model, while the use of the difference corresponds to an additive model (Rothman et al., 1980; Saracci, 1980). A distinction is sometimes made between 'statistical interaction' which can be a departure from either a multiplicative or additive model, and 'biologic interaction' which is measured by departure from an additive model (Rothman, 2012). However, neither additive nor multiplicative models point to a particular biologic mechanism. Regardless of the model choice, the main objective is to understand how the joint effect of two exposures differs from their separate effects (in the absence of the other exposure). The Human Genomic Epidemiology Network (HuGENet) proposed a layout for transparent presentation of separate and joint effects that permits evaluation of different types of interaction (Botto and Khoury, 2001). A difficulty is that some study designs, such as case–control studies, and several statistical models, such as logistic or Cox regression models, estimate risk or rate ratios and intrinsically lead to multiplicative modelling.

the formulae used or a citation for the formula (Uter and Pfahlberg, 2001).

17. Other analyses: report other analyses done, such as sensitivity / robustness analysis and analysis of subgroups

Examples

Sensitivity testing was done for each model by lowering prior test accuracy estimates' mode by 10 percentage points (e.g., $ELISA_{Se}$ from 0.95 to 0.85), relaxing the beta distribution to a 50^{th} percentile and reducing the lower bound 10 percentage points below the previous lower bound (e.g., $ELISA_{Se}$ from 0.9 to 0.8) and using a uniform beta distribution as the prior distribution for exposure prevalence (Haley et al., 2011).

Explanation

When an observational study has a single primary question, the reader reasonably might assume that all the study design features were selected with that question in mind (e.g. sample size and power, the interpretation of the alpha error, accuracy enhancement and bias reduction measures and potential confounders). If additional questions and analyses were included in the study, the authors must tell the readers. Lack of full disclosure distorts the interpretation of everything from bias control effectiveness to multiplicity considerations. The reader must be informed of all secondary analyses (e.g. conducting sensitivity analyses, or testing for interactions or particular subgroup analyses) were pre-specified (*a priori*, e.g.) or were steered by the data themselves (*post hoc* analyses). For example, 'non-significant' interactions or risk factors are 'results' unto themselves, and they assist in framing the context of 'significant' results. *Post hoc* subgroup analyses that appear more 'exciting' than the answers to the primary question must be viewed cautiously.

Researchers often must extrapolate the 'base' values of relevant input data. Examples include declaring prior distributions for Bayesian analyses and declaring the diagnostic test accuracy used for adjusting apparent prevalences to true prevalences. In some instances, distributions are unknown or variables have great inherent contextual variation, which leads to considerable uncertainty. When faced with such assumptions about uncertainty, authors should conduct 'sensitivity analyses' to discover the sensitivity (robustness) of the conclusions with respect to reasonable variation from base values. Additionally, the data often lead to decisions about diagnoses or other categorizations; however, alternative decision criteria might be available. In these situations, it is also appropriate to examine and report the sensitivity of the conclusions with respect to the decision criteria used.

Discussion

The discussion section provides readers with the authors' interpretation of the results once they have been placed in context such as the approach to the study and prior relevant findings. Authors should also emphasize the study design aspects that enhance the internal and external validity of the findings to aid the readers' understanding of the data and the conclusion's strength. In addition, the authors should outline the limitations of the design and their impact on the findings.

18. Summarize key results with reference to study objectives

Examples

. . .a large field data collection from 14 endemically infected dairy herds was used to investigate the hypothesis that cattle with persistently high antibody levels are at high risk of shedding *S.* Dublin and therefore are candidates to be culled or at least managed so that they do not spread the

infection to herd mates. Despite that fact that there were seropositive animals in many of the age groups at most of the herd visits, indicative of the herds being endemically infected, the general probability of shedding was very low . . . for *S*. Dublin. . . . Based on this study material there was no evidence that animals with persistently high antibodies over a period of at least 6 months were at higher risk of shedding *S*. Dublin bacteria in their faeces than other seropositive cattle (Nielsen, 2013).

Explanation

It is good practice to begin the discussion with a short summary of the main findings of the study. The short summary reminds readers of the main findings and might help them assess whether the subsequent interpretation and implications offered by the authors are supported by the findings.

19. Discuss strengths and limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias

Example

Because of the steps involved in making a diagnosis of leptospirosis . . . it is unlikely that a dog examined at a veterinary teaching hospital would be falsely diagnosed as having leptospirosis when it did not. It is more likely that leptospirosis was not diagnosed in some dogs with the disease. This of bias is unlikely to be substantial, because the number of dogs with undiagnosed leptospirosis is probably a very small proportion of all dogs examined at veterinary teaching hospitals. For example, if leptospirosis had been underdiagnosed by a factor of 10-fold, < 0.04% of the 1,819,792 dogs examined at veterinary teaching hospitals between 1970 and 1998 would have been classified as cases rather than controls. For the age category 4 to 6.9 years, this bias would result in a change in the estimated OR from 1.7259 to 1.7295 (a change of 0.21%), assuming equal proportions of misdiagnoses in the 4 to 6.9 years and < 1 year age categories. We do not expect the proportion of dogs with leptospirosis in which the disease is not diagnosed at veterinary teaching hospitals to be greater than 10-fold that recorded in the VMDB, so bias from misclassification of leptospirosis status was unlikely to be substantial in this study (Ward et al., 2002).

Explanation

Authors should highlight specific strengths of their study relative to other work in the field (e.g. a study based on true random sampling versus convenience sampling). However, the identification and discussion of the limitations of a study are an essential part of scientific reporting. It is important not only to identify the sources of bias and confounding that could have affected results, but also to discuss the relative importance of different biases, including the likely direction and magnitude of any potential bias (see Box 1: Bias in observational studies and Box 5: Confounding and item 9 about bias in method and materials).

Authors should also discuss the impact of imprecision and uncertainty on the interpretation of results. Result imprecision could result from a small sample size, which produces a wide CI such as low effect size precision. Here, we refer to uncertainty as missing knowledge related to specific factors, parameters or model specification rather than sample size (Firestone et al., 1997).

When discussing limitations, authors might compare the study being presented with other studies in the literature in terms of validity, generalizability and precision. In this approach, each study can be viewed as a contribution to the literature, not as a stand-alone basis for inference and action (Poole et al., 2003).

20. Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies and other relevant evidence

Examples

We conclude that the presence of unresolved infection in a herd is a contributor to further bTB episodes in the first 2 years after clearance. These findings agree with the investigations in the UK and Ireland, which have shown repeatedly that bTB spreads from de-restricted herds to clear herds via the transfer of undetected infection after de-restriction [] (Dawson et al., 2014).

Explanation

In accordance with the original STROBE document, we encourage authors to provide the reader with a thoughtful conclusion and a rationale based on the principles of causal inference rather than using *P*-values less than 0.05 (or any other arbitrary *P*-value cut-off) as an indicator of a causal association. The heart of the discussion section is the interpretation of a study's results. When interpreting results, authors should consider the place of the study on the discovery-to-verification continuum and potential sources of bias, including loss to follow-up and non-participation (see also items 9, 12 and 19).

In the veterinary field, studies evaluating large numbers of independent variables are common occurrences. The probability that at least one significant finding will be a type I error increases as the number of hypotheses tested within a study increases. Therefore, in the limitations section, authors should note the probability of type I errors as an

alternative explanation for the associations observed when appropriate.

The rationale should address the concepts used to establish causation. The conclusions presented should consider the role chance and bias could play in the findings of the current study (discussed in item 19) as well as those of previous studies on the same topic. Currently, many guides on causal thinking exist, such as those proposed by Bradford Hill and others (Kaufman and Poole, 2000; Rothman and Greenland, 2005; Hill, 2015). While we are not proposing a formulaic application of guidelines or criteria, readers might find it helpful if the authors document the concepts of causal inference to assist them in understanding the conclusion. For example, how strong is the association with the exposure? Did exposure precede disease onset? Is the association consistently observed in different studies and settings? Is there supporting evidence from experimental studies, including laboratory and animal studies? How specific is the exposure's putative effect, and is there a dose–response relationship? Is the association biologically plausible? A discussion of the existing external evidence, from different types of studies, should always be included, but might be particularly important for studies reporting small increases in risk. Furthermore, authors should put their results in context with similar studies and explain how the new study affects the existing body of evidence, ideally by referring to a systematic review.

21. Discuss the generalizability (external validity) of the study results

Example

The findings from this study would be difficult to extrapolate to other countries, because of the differences in bTB management policies between countries. However, this study has added weight to the growing body of evidence to show that residual infection in herds poses a problem to bTB eradication schemes, and that the goal should be to maximize within-herd sensitivity in the management of this problematic infection (Dawson et al., 2014).

Explanation

Generalizability, also called external validity or applicability, is the extent to which the results of a study can be applied to other circumstances (Campbell, 1957). There is no external validity *per se*; the term is meaningful only with regard to clearly specified conditions (Justice et al., 1999). Can results be applied to an individual, groups or populations that differ from those enrolled in the study with regard to age, sex, breed or other characteristic, such as the production system for livestock populations? Are the nature and level of exposures comparable, and the definitions of outcomes relevant to another setting or population? Are results from one country applicable to other countries?

The question of whether the results of a study have external validity is often a matter of judgment that depends on the study setting, the characteristics of the participants, the exposures examined and the outcomes assessed. Thus, it is crucial that authors provide readers with adequate information about the setting and locations, eligibility criteria, the exposures and how they were measured, the definition of outcomes and the period of recruitment and follow-up. The degree of non-participation and the proportion of unexposed participants in whom the outcome develops are also relevant. Knowledge of the absolute risk and prevalence of the exposure, which will often vary across populations, are helpful when applying results to other settings and populations. Of course, the need for inclusion of these features is discussed throughout this document and these summary statements only serve as a reminder of the need for complete reporting of research design and results in the context of external validity (See Box 6: Measures of Association and measures of impact.).

Other information

22. Transparency

22 (a). Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

Example

Funding: This study was funded by Pfizer Animal Health (www.Zoetis.com). The grant number was 1329. The following Pfizer personnel were observers during the study: Jeremy Salt, Michael Pearce, Tony Simon and Marie-Odile Hendrickx. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript, except attendance at project coordination meetings as observers, and commenting on the draft manuscript (Jones et al., 2013).

Explanation

Because of concerns about funding agencies' influence on study design and the potential for selective reporting, funding sources and the role of funding agencies should be described explicitly. In human health, several investigations show strong associations between the source of funding and the conclusions of research articles (Davidson, 1986; Stelfox et al., 1998; Bekelman et al., 2003; Lexchin et al., 2003). To our knowledge, there are no similar associations detected in observational studies conducted in veterinary science. However, the possibility for bias could exist, and it is best practice to disclose all funding sources.

22 (b). Conflict of interests. Describe any conflict of interests, or lack thereof, for each author

Dr. . . . has provided scientific consulting services to Zoetis Inc. (manufacturer of the CCFA product). This does not alter the authors' adherence to all the journal policies on sharing data and materials (Kanwar et al., 2014)

Explanation

In human health, there is evidence that authors or funders might have conflict of interests that influence any of the following: the design of the study (Safer, 2002), choice of exposures (Aspinall and Goodman, 1995; Safer, 2002), outcomes (Chan et al., 2004), statistical methods (Melander et al., 2003) and selective publication of outcomes and studies (Chan et al., 2004; Scherer et al., 2007). Potential conflict of interests include financial arrangements outside of research funding that could influence authors. Authors should disclose any financial support, including grants, scholarships and sponsorships received. Gifts that might not be associated directly with the project, such as laboratory equipment, travel, consulting agreements and honoraria, but still establish a relationship with a company or agency should also be disclosed. This information alerts users the relationship and allows them to assess the potential for bias in conducting and reporting the study.

22 (c). Describe the authors' roles – provision of an authors' declaration of transparency is recommended

Example

H.M.S., B.N., G.H.L. and P.B. conceived and designed the study. N.K., H.M.S., B.N., G.H.L., J.V., P.B., J.L.C. and G.C. performed the experiments. N.K., H.M.S., B.N., G.H.L., P.B., M.M.C. and J.B. analyzed and interpreted data. N.K. and H.M.S. drafted the manuscript. All authors revised manuscript for critically important intellectual content and approved the final version to be published (Kanwar et al., 2014).

Explanation

A declaration of transparency should contain two parts: the authors' roles in the study and a declaration of complete reporting. Declarations of transparency do not address any potential bias within the study. They are primarily designed to ensure that all authors meet the criteria for authorship.[3] Those persons not meeting the authorship criteria can be referenced in the acknowledgements section, but it is unethical to include them as authors. The declaration of complete reporting was proposed by Altman (Altman and Moher, 2013) and is a statement acknowledging that the data have been reported in their entirety and none of the study has been omitted, thus reducing the potential for selective reporting.

[3]http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html

22 (d). Ethical approval – include information on ethical approval for use of animal and human subjects

Example

The study was performed in adherence to the University of Liverpool Animal Ethics Guidelines (Trehy et al., 2014).

Explanation

When studies involve animal use, authors must obtain ethical approval. It is consistent with best practices and transparency to report the agency in charge of approval and a verification number to identify the approval. In some studies, particularly those using questionnaires, authors must document that they received approval for recruiting human participants in research.

22 (e). Quality standards – describe any quality standards used in the conduct of the research

Example

. . .reported according to the guidelines of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement, see additional information (STROBE Checklist) for further details (Forand, 2004).

Explanation

When standards, such as STROBE, are available and have been used, authors should explicitly state their use, so end-users are aware of their implementation and can validate that they were followed. Further, some standards are lengthy, and it might be impractical to include a full description of the methods employed in the paper. For example, laboratory or animal welfare accreditations indicate that certain practices and quality control approaches were followed, and this information can help end-users assess bias.

References

Ahlman, T., B. Berglund, L. Rydhmer, and E. Strandberg, 2011: Culling reasons in organic and conventional dairy herds and genotype by environment interaction for longevity. *J. Dairy Sci.* 94, 1568–1575.

Alawneh, J. I., T. S. Barnes, C. Parke, E. Lapuz, E. David, V. Basinang, A. Baluyut, E. Villar, E. L. Lopez, and P. J. Blackall, 2014: Description of the pig production systems, biosecurity

practices and herd health providers in two provinces with high swine density in the Philippines. *Prev. Vet. Med.* 114, 73–87.

Al-Shahi Salman, R., E. Beller, J. Kagan, E. Hemminki, R. S. Phillips, J. Savulescu, M. Macleod, J. Wisely, and I. Chalmers, 2014: Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 383, 176–185.

Altman, D. G., 2005: Categorizing continuous variables. In: Armitage, P., and T Colton (eds.), Encyclopedia of Biostatistics, John Wiley & Sons Ltd, Hoboken, NJ.

Altman, D. G., and D. Moher, 2013: Declaration of transparency for each research article. *BMJ* 347, f4796.

Altman, D. G., B. Lausen, W. Sauerbrei, and M. Schumacher, 1994: Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. *J. Natl Cancer Inst.* 86, 829–835.

Anderson, H. R., R. W. Atkinson, J. L. Peacock, M. J. Sweeting, and L. Marston, 2005: Ambient particulate matter and health effects: publication bias in studies of short-term associations. *Epidemiology* 16, 155–163.

Anonymous, 2012: Principles of Epidemiology in Public Health Practice, Third Edition. An Introduction to Applied Epidemiology and Biostatistics. Centers for Disease Control, Atlanta, Georgia.

Aspinall, R. L., and N. W. Goodman, 1995: Denial of effective treatment and poor quality of clinical information in placebo controlled trials of ondansetron for postoperative nausea and vomiting: a review of published trials. *BMJ* 311, 844–846.

Babcock, A. H., N. Cernicchiaro, B. J. White, S. R. Dubnicka, D. U. Thomson, S. E. Ives, H. M. Scott, G. A. Milliken, and D. G. Renter, 2013: A multivariable assessment quantifying effects of cohort-level factors associated with combined mortality and culling risk in cohorts of U.S. commercial feedlot cattle. *Prev. Vet. Med.* 108, 38–46.

Bamaiyi, P. H., L. Hassan, S. Khairani-Bejo, M. ZainalAbidin, M. Ramlan, N. Krishnan, A. Adzhar, N. Abdullah, N. H. Hamidah, M. M. Norsuhanna, and S. N. Hashim, 2014: Case-control study on risk factors associated with Brucella melitensis in goat farms in Peninsular Malaysia. *Trop. Anim. Health Prod.* 46, 739–745.

Becher, H., 1992: The concept of residual confounding in regression models and some applications. *Stat. Med.* 11, 1747–1758.

Bekelman, J. E., Y. Li, and C. P. Gross, 2003: Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *J. Am. Med. Assoc.* 289, 454–465.

Benson, K., and A. J. Hartz, 2000: A comparison of observational studies and randomized, controlled trials. *N. Engl. J. Med.* 342, 1878–1886.

Botto, L. D., and M. J. Khoury, 2001: Commentary: facing the challenge of gene-environment interaction: the two-by-four table and beyond. *Am. J. Epidemiol.* 153, 1016–1020.

Brenner, H., and M. Blettner, 1997: Controlling for continuous confounders in epidemiologic research. *Epidemiology* 8, 429–434.

Burow, E., P. T. Thomsen, T. Rousing, and J. T. Sørensen, 2014: Track way distance and cover as risk factors for lameness in Danish dairy cows. *Prev. Vet. Med.* 113, 625–628.

Campbell, D. T., 1957: Factors relevant to the validity of experiments in social settings. *Psychol. Bull.* 54, 297–312.

Carlin, J. B., and L. W. Doyle, 2002: Sample size. *J. Paediatr. Child Health* 38, 300–304.

Cave, N. J., F. J. Allan, S. L. Schokkenbroek, C. A. Metekohy, and D. U. Pfeiffer, 2012: A cross-sectional study to compare changes in the prevalence and risk factors for feline obesity between 1993 and 2007 in New Zealand. *Prev. Vet. Med.* 107, 121–133.

Chan, A. W., A. Hróbjartsson, M. T. Haahr, P. C. Gøtzsche, and D. G. Altman, 2004: Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *J. Am. Med. Assoc.* 291, 2457–2465.

Chan, A. W., F. Song, A. Vickers, T. Jefferson, K. Dickersin, P. C. Gotzsche, H. M. Krumholz, D. Ghersi, and H. B. van der Worp, 2014: Increasing value and reducing waste: addressing inaccessible research. *Lancet* 383, 257–266.

Christenfeld, N. J., R. P. Sloan, D. Carroll, and S. Greenland, 2004: Risk factors, confounding, and the illusion of statistical control. *Psychosom. Med.* 66, 868–875.

Clark, T. G., D. G. Altman, and B. L. De Stavola, 2002: Quantification of the completeness of follow-up. *Lancet* 359, 1309–1310.

Clayton, D., M. Hills, and A. Pickles, 1993: Statistical Models in Epidemiology. Oxford University Press, Oxford.

Cochran, W. G., 1968: The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics* 24, 295–313.

Cohen, J., 1983: The cost of dichotomization.

Cole, P., and B. MacMahon, 1971: Attributable risk percent in case-control studies. *Br. J. Prev. Soc. Med.* 25, 242–244.

Costanza, M. C., 1995: Matching. *Prev. Med.* 24, 425–433.

Cox, D. R., 1957: Note on grouping. *J. Am. Stat. Assoc.* 52, 543–547.

Custer, B., W. T. Jr Longstreth, L. E. Phillips, T. D. Koepsell, and G. Van Belle, 2006: Hormonal exposures and the risk of intracranial meningioma in women: a population-based case-control study. *BMC Cancer* 6, 152.

Davidson, R. A., 1986: Source of funding and outcome of clinical trials. *J. Gen. Intern. Med.* 1, 155–158.

Dawson, K. L., M. A. Stevenson, J. A. Sinclair, and M. A. Bosson, 2014: Recurrent bovine tuberculosis in New Zealand cattle and deer herds 2006–2010. *Epidemiol. Infect.* 142, 2065–2074.

De Vliegher, S., H. Laevens, H. W. Barkema, I. R. Dohoo, H. Stryhn, G. Opsomer, and A. de Kruif, 2004: Management practices and heifer characteristics associated with early lactation somatic cell count of Belgian dairy heifers. *J. Dairy Sci.* 87, 937–947.

Dean, R. S., D. U. Pfeiffer, and V. J. Adams, 2013: The incidence of feline injection site sarcomas in the United Kingdom. *BMC Vet. Res.* 9, 17.

Devine, O. J., and J. M. Smith, 1998: Estimating sample size for epidemiologic studies: the impact of ignoring exposure measurement uncertainty. *Stat. Med.* 17, 1375–1389.

Dewell, R. D., S. T. Millman, S. A. Gould, K. L. Tofflemire, R. D. Whitley, R. L. Parsons, E. W. Rowe, F. Liu, C. Wang, and A. M. O'Connor, 2014: Evaluating approaches to measuring ocular pain in bovine calves with corneal scarification and infectious bovine keratoconjunctivitis-associated corneal ulcerations. *J. Anim. Sci.* 92, 1161–1172.

Dhand, N. K., J. Eppleston, R. J. Whittington, and J. A. Toribio, 2007: Risk factors for ovine Johne's disease in infected sheep flocks in Australia. *Prev. Vet. Med.* 82, 51–71.

Diaz-Ordaz, K., M. G. Kenward, A. Cohen, C. L. Coleman, and S. Eldridge, 2014: Are missing data adequately handled in cluster randomised trials? A systematic review and guidelines. *Clin. Trials* 11, 590–600.

Dohoo, I. R., 2014: Bias–is it a problem, and what should we do? *Prev. Vet. Med.* 113, 331–337.

Dohoo, I. R., 2015: Dealing with deficient and missing data. *Prev. Vet. Med.* 122, 221–228.

Dohoo, I. R., E. Tillard, H. Stryhn, and B. Faye, 2001: The use of multilevel models to evaluate sources of variation in reproductive performance in dairy cattle in Reunion Island. *Prev. Vet. Med.* 50, 127–144.

Dohoo, I. R., S. W. Martin, and H. Stryhn, 2009: Veterinary Epidemiologic Research. VER Incorporated, Charlottetown, PEI, Canada.

Donders, A. R., G. J. van der Heijden, T. Stijnen, and K. G. Moons, 2006: Review: a gentle introduction to imputation of missing values. *J. Clin. Epidemiol.* 59, 1087–1091.

Drescher, K., J. Timm, and K. H. Jockel, 1990: The design of case-control studies: the effect of confounding on sample size requirements. *Stat. Med.* 9, 765–776.

Dufour, S., H. W. Barkema, L. DesCôteaux, T. J. DeVries, I. R. Dohoo, K. Reyher, J.-P. Roy, and D. T. Scholl, 2010: Development and validation of a bilingual questionnaire for measuring udder health related management practices on dairy farms. *Prev. Vet. Med.* 95, 74–85.

Dunn, N., A. Arscott, and M. Thorogood, 2001: The relationship between use of oral contraceptives and myocardial infarction in young women with fatal outcome, compared to those who survive: results from the MICA case-control study. *Contraception* 63, 65–69.

Egger, M., P. Jüni, C. Bartlett, and C. G. C. S. o. R. o. Trials, 2001: Value of flow diagrams in reports of randomized controlled trials. *J. Am. Med. Assoc.* 285, 1996–1999.

von Elm, E., D. G. Altman, M. Egger, S. J. Pocock, P. C. Gotzsche, J. P. Vandenbroucke, and S. Initiative, 2007: Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 335, 806–808.

Emanuelson, U., and A. Egenvall, 2014: The data - sources and validation. *Prev. Vet. Med.* 113, 298–303.

European Food Safety Authority, 2013: Sample size considerations for hierarchical population. *EFSA J.* 11, 47.

Evans, D., B. Chaix, T. Lobbedez, C. Verger, and A. Flahault, 2012: Combining directed acyclic graphs and the change-in-estimate procedure as a novel approach to adjustment-variable selection in epidemiology. *BMC Med. Res. Methodol.* 12, 156.

Fanelli, D., 2013: Redefine misconduct as distorted reporting. *Nature* 494, 149.

Fast, R., T. Schütt, N. Toft, A. Møller, and M. Berendt, 2013: An observational study with long-term follow-up of canine cognitive dysfunction: clinical characteristics, survival, and risk factors. *J. Vet. Intern. Med.* 27, 822–829.

Fediaevsky, A., E. Morignat, C. Ducrot, and D. Calavas, 2009: A case-control study on the origin of atypical scrapie in sheep, France. *Emerg. Infect. Dis.* 15, 710–718.

Firestone, M., P. Fenner-Crisp, T. Barry, D. Bennett, S. Chang, M. Callahan, A. Burke, J. Michaud, M. Olsen, and P. Cirone, 1997: Guiding principles for Monte Carlo analysis. US Environmental Protection Agency, Washington, DC.

Forand, S. P., 2004: Leukaemia incidence among workers in the shoe and boot manufacturing industry: a case-control study. *Environ. Health* 3, 7.

Funk, L. D., J. M. Reecy, C. Wang, R. G. Jr Tait, and A. M. O'Connor, 2014: Associations between infectious bovine keratoconjunctivitis at weaning and ultrasongraphically measured body composition traits in yearling cattle. *J. Am. Vet. Med. Assoc.* 244, 100–106.

Galea, S., and M. Tracy, 2007: Participation rates in epidemiologic studies. *Ann. Epidemiol.* 17, 643–653.

Gefeller, O., A. Pfahlberg, H. Brenner, and J. Windeler, 1998: An empirical investigation on matching in published case-control studies. *Eur. J. Epidemiol.* 14, 321–325.

Giuffrida, M. A., K. A. Agnello, and D. C. Brown, 2012: Blinding terminology used in reports of randomized controlled trials involving dogs and cats. *J. Am. Vet. Med. Assoc.* 241, 1221–1226.

Greenland, S., 1995: Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. *Epidemiology* 6, 450–454.

Greenland, S., 1998: Applications of stratified analysis methods. In: Rothman K. J., and S. Greenland (eds), Modern Epidemiology, Lippincott Raven, Philadelphia, PA.

Greenland, S., 2003: The impact of prior distributions for uncontrolled confounding and response bias: a case study of the relation of wire codes and magnetic fields to childhood leukemia. *J. Am. Stat. Assoc.* 98, 47–54.

Greenland, S., and B. Brumback, 2002: An overview of relations among causal modelling methods. *Int. J. Epidemiol.* 31, 1030–1037.

Greenland, P., M. L. Daviglus, A. R. Dyer, K. Liu, C. F. Huang, J. J. Goldberger, and J. Stamler, 1999: Resting heart rate is a risk factor for cardiovascular and noncardiovascular mortality: the Chicago Heart Association Detection Project in Industry. *Am. J. Epidemiol.* 149, 853–862.

Guerin, M. T., W. Martin, J. Reiersen, O. Berke, S. A. McEwen, J.-R. Bisaillon, and R. Lowman, 2007: A farm-level study of risk factors associated with the colonization of broiler flocks with Campylobacter spp. in Iceland, 2001–2004. *Acta Vet. Scand.* 49, 18.

Hailu, B., T. Tolosa, G. Gari, T. Teklue, and B. Beyene, 2014: Estimated prevalence and risk factors associated with clinical Lumpy skin disease in north-eastern Ethiopia. *Prev. Vet. Med.* 115, 64–68.

Haley, C., B. Wagner, S. Puvanendiran, J. Abrahante, and M. P. Murtaugh, 2011: Diagnostic performance measures of ELISA and quantitative PCR tests for porcine circovirus type 2 exposure using Bayesian latent class analysis. *Prev. Vet. Med.* 101, 79–88.

Hay, K. E., T. S. Barnes, J. M. Morton, A. C. Clements, and T. J. Mahony, 2014: Risk factors for bovine respiratory disease in Australian feedlot cattle: use of a causal diagram-informed approach to estimate effects of animal mixing and movements before feedlot entry. *Prev. Vet. Med.* 117, 160–169.

Haynes, A. B., W. R. Berry, and A. A. Gawande, 2015: What do we know about the safe surgery checklist now? *Ann. Surg.* 261, 829–830.

Hernan, M. A., S. Hernandez-Diaz, M. M. Werler, and A. A. Mitchell, 2002: Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am. J. Epidemiol.* 155, 176–184.

Hernán, M. A., S. Hernández-Díaz, and J. M. Robins, 2004: A structural approach to selection bias. *Epidemiology* 15, 615–625.

Hill, A. B., 2015: The environment and disease: association or causation? *J. R. Soc. Med.* 108, 32–37.

Holm, D. E., M. Nielen, R. Jorritsma, P. C. Irons, and P. N. Thompson, 2015: Evaluation of pre-breeding reproductive tract scoring as a predictor of long term reproductive performance in beef heifers. *Prev. Vet. Med.* 118, 56–63.

Il'yasova , D., I. Hertz-Picciotto, U. Peters, J. A. Berlin, and C. Poole, 2005: Choice of exposure scores for categorical regression in meta-analysis: a case study of a common problem. *Cancer Causes Control*, 16, 383–388.

Ioannidis, J. P., S. Greenland, M. A. Hlatky, M. J. Khoury, M. R. Macleod, D. Moher, K. F. Schulz, and R. Tibshirani, 2014: Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 383, 166–175.

Jones, B. A., C. Sauter-Louis, J. Henning, A. Stoll, M. Nielen, G. Van Schaik, A. Smolenaars, M. Schouten, I. den Uijl, C. Fourichon, R. Guatteo, A. Madouasse, S. Nusinovici, P. Deprez, S. De Vliegher, J. Laureyns, R. Booth, J. M. Cardwell, and D. U. Pfeiffer, 2013: Calf-level factors associated with bovine neonatal pancytopenia–a multi-country case-control study. *PLoS ONE* 8, e80619.

Justice, A. C., K. E. Covinsky, and J. A. Berlin, 1999: Assessing the generalizability of prognostic information. *Ann. Intern. Med.* 130, 515–524.

Kadohira, M., J. J. McDermott, M. M. Shoukri, and M. A. Thorburn, 1997: Assessing infections at multiple levels of aggregation. *Prev. Vet. Med.* 29, 161–177.

Kaneene, J. B., C. S. Bruning-Fann, L. M. Granger, R. Miller, and B. A. Porter-Spalding, 2002: Environmental and farm management factors associated with tuberculosis on cattle farms in northeastern Michigan. *J. Am. Vet. Med. Assoc.* 221, 837–842.

Kanwar, N., H. M. Scott, B. Norby, G. H. Loneragan, J. Vinasco, J. L. Cottell, G. Chalmers, M. M. Chengappa, J. Bai, and P. Boerlin, 2014: Impact of treatment strategies on cephalosporin and tetracycline resistance gene quantities in the bovine fecal metagenome. *Sci. Rep.* 4, 5100.

Kaufman, J. S., and C. Poole, 2000: Looking back on "causal thinking in the health sciences". *Annu. Rev. Public Health* 21, 101–119.

KilBride, A. L., M. Mendl, P. Statham, S. Held, M. Harris, S. Cooper, and L. E. Green, 2012: A cohort study of preweaning piglet mortality and farrowing accommodation on 112 commercial pig farms in England. *Prev. Vet. Med.* 104, 281–291.

Kim, R. Y., G. Kwakye, A. C. Kwok, R. Baltaga, G. Ciobanu, A. F. Merry, L. M. Funk, S. R. Lipsitz, A. A. Gawande, W. R. Berry, and A. B. Haynes, 2015: Sustainability and long-term effectiveness of the WHO surgical safety checklist combined with pulse oximetry in a resource-limited setting: two-year update from Moldova. *JAMA Surg.* 150, 473–479.

Kleinbaum, D., 2009: ActivEpi. Springer, New York.

Kristiansen, V. M., A. Nodtvedt, A. M. Breen, M. Langeland, J. Teige, M. Goldschmidt, T. J. Jonasdottir, T. Grotmol, and K. Sorenmo, 2013: Effect of ovariohysterectomy at the time of tumor removal in dogs with benign mammary tumors and hyperplastic lesions: a randomized controlled clinical trial. *J. Vet. Intern. Med.* 27, 935–942.

Lang, T. A., and D. G. Altman, 2013: *Basic Statistical Reporting for Articles Published in Biomedical Journals: The "Statistical Analyses and Methods in the Published Literature" or The SAMPL Guidelines*".

Lash, T. L., and A. K. Fink, 2003: Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology* 14, 451–458.

Lehner, G., M. Linek, R. Bond, D. H. Lloyd, E. Prenger-Berninghoff, N. Thom, I. Straube, K. Verheyen, and A. Loeffler, 2014: Case-control risk factor study of methicillin-resistant Staphylococcus pseudintermedius (MRSP) infection in dogs and cats in Germany. *Vet. Microbiol.* 168, 154–160.

Lexchin, J., L. A. Bero, B. Djulbegovic, and O. Clark, 2003: Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 326, 1167–1170.

Lipsitz, S. R., J. G. Ibrahim, M. H. Chen, and H. Peterson, 1999: Non-ignorable missing covariates in generalized linear models. *Stat. Med.* 18, 2435–2448.

Little, R. J. A., and D. B. Rubin, 2002: Statistical Analysis with Missing Data. Wiley, Hoboken, NJ.

Lupo, C., A. Osta Amigo, Y. V. Mandard, C. Peroz, and T. Renault, 2014: Improving early detection of exotic or emergent oyster diseases in France: Identifying factors associated with shellfish farmer reporting behaviour of oyster mortality. *Prev. Vet. Med.* 116, 168–182.

MacCallum, R. C., S. Zhang, K. J. Preacher, and D. D. Rucker, 2002: On the practice of dichotomization of quantitative variables. *Psychol. Methods* 7, 19–40.

Macleod, M. R., S. Michie, I. Roberts, U. Dirnagl, I. Chalmers, J. P. Ioannidis, R. Al-Shahi Salman, A. W. Chan, and P. Glasziou, 2014: Biomedical research: increasing value, reducing waste. *Lancet* 383, 101–104.

Maddox, T. W., P. D. Clegg, P. J. Diggle, A. L. Wedley, S. Dawson, G. L. Pinchbeck, and N. J. Williams, 2012: Cross-sectional study of antimicrobial-resistant bacteria in horses. Part 1: Prevalence of antimicrobial-resistant Escherichia coli and methicillin-resistant Staphylococcus aureus. *Equine Vet. J.* 44, 289–296.

Madsen, J. M., N. G. Zimmermann, J. Timmons, and N. L. Tablante, 2013: Avian influenza seroprevalence and biosecurity risk factors in Maryland backyard poultry: a cross-sectional study. *PLoS ONE* 8, e56851.

McDermott, J. J., M. Kadohira, C. J. O'Callaghan, and M. M. Shoukri, 1997: A comparison of different models for assessing variations in the sero-prevalence of infectious bovine rhinotracheitis by farm, area and district in Kenya. *Prev. Vet. Med.* 32, 219–234.

McDowell, S. W. J., F. D. Menzies, S. H. McBride, A. N. Oza, J. P. McKenna, A. W. Gordon, and S. D. Neill, 2008: Campylobacter spp. in conventional broiler flocks in Northern Ireland: epidemiology and risk factors. *Prev. Vet. Med.* 84, 261–276.

Melander, H., J. Ahlqvist-Rastad, G. Meijer, and B. Beermann, 2003: Evidence b(i)ased medicine–selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 326, 1171–1173.

Menzies, F. D., D. A. Abernethy, L. A. Stringer, N. Honhold, and A. W. Gordon, 2012: A matched cohort study investigating the risk of Mycobacterium bovis infection in the progeny of infected cows. *Vet. J.* 194, 299–302.

Miettinen, O. S., 1985: Theoretical Epidemiology: Principles of Occurrence Research in Medicine. Wiley, New York.

Moher, D., K. F. Schulz, and D. G. Altman, 2001: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 357, 1191–1194.

Moher, D., P. Glasziou, I. Chalmers, M. Nasser, P. M. Bossuyt, D. A. Korevaar, I. D. Graham, P. Ravaud, and I. Boutron, 2016: Increasing value and reducing waste in biomedical research: who's listening? *Lancet* 387, 1573–1586.

Murray, J. K., E. R. Singer, K. L. Morgan, C. J. Proudman, and N. P. French, 2004: Memory decay and performance-related information bias in the reporting of scores by event riders. *Prev. Vet. Med.* 63, 173–182.

Nielsen, L. R., 2013: Salmonella Dublin faecal excretion probabilities in cattle with different temporal antibody profiles in 14 endemically infected dairy herds. *Epidemiol. Infect.* 141, 1937–1944.

Nielsen, L. R., L. D. Warnick, and M. Greiner, 2007: Risk factors for changing test classification in the Danish surveillance program for Salmonella in dairy herds. *J. Dairy Sci.* 90, 2815–2825.

Nødtvedt, A., I. Dohoo, J. Sanchez, G. Conboy, L. DesCȷ̂teaux, G. Keefe, K. Leslie, and J. Campbell, 2002: The use of negative binomial modelling in a longitudinal study of gastrointestinal parasite burdens in Canadian dairy cows. *Can. J. Vet. Res.* 66, 249–257.

O'Connor, A. M., J. M. Sargeant, I. A. Gardner, J. S. Dickson, M. E. Torrence, C. E. Dewey, I. R. Dohoo, R. B. Evans, J. T. Gray, M. Greiner, G. Keefe, S. L. Lefebvre, P. S. Morley, A. Ramirez, W. Sischo, D. R. Smith, K. Snedeker, J. Sofos, M. P. Ward, and R. Wills, 2010: The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety. *J. Vet. Intern. Med.* 24, 57–64.

Okafor, C. C., S. L. Lefebvre, D. L. Pearl, M. Yang, M. Wang, S. L. Blois, E. M. Lund, and C. E. Dewey, 2014: Risk factors associated with calcium oxalate urolithiasis in dogs evaluated at general care veterinary hospitals in the United States. *Prev. Vet. Med.* 115, 217–228.

Olea-Popelka, F. J., J. Phelan, P. W. White, G. McGrath, J. D. Collins, J. O'Keeffe, M. Duggan, D. M. Collins, D. F. Kelton, O. Berke, S. J. More, and S. W. Martin, 2006: Quantifying badger exposure and the risk of bovine tuberculosis for cattle herds in county Kilkenny, Ireland. *Prev. Vet. Med.* 75, 34–46.

Olsen, J., and O. Basso, 1999: Re: Residual confounding. *Am. J. Epidemiol.* 149, 290.

Pearce, N., 2012: Classification of epidemiological study designs. *Int. J. Epidemiol.* 41, 393–397.

Phillips, C. V., 2003: Quantifying and reporting uncertainty from systematic errors. *Epidemiology* 14, 459–466.

Pillars, R. B., M. W. Bolton, and D. L. Grooms, 2011: Case-control study: productivity and longevity of dairy cows that tested positive for infection with Mycobacterium avium ssp. paratuberculosis as heifers compared to age-matched controls. *J. Dairy Sci.* 94, 2825–2831.

Pires, A. F., J. A. Funk, R. Manuzon, M. Darr, and L. Zhao, 2013: Longitudinal study to evaluate the association between thermal environment and Salmonella shedding in a midwestern US swine farm. *Prev. Vet. Med.* 112, 128–137.

Poole, C., U. Peters, D. Il'yasova, and L. Arab, 2003: Commentary: this study failed? *Int. J. Epidemiol.* 32, 534–535.

Ramon, M. E., M. R. Slater, M. P. Ward, and R. R. Lopez, 2008: Repeatability of a telephone questionnaire on cat-ownership patterns and pet-owner demographics evaluation in a community in Texas, USA. *Prev. Vet. Med.* 85, 23–33.

Rigby, A. S., and A. Vail, 1998: Statistical methods in epidemiology. II: a commonsense approach to sample size estimation. *Disabil. Rehabil.* 20, 405–410.

Rizzo, F., N. Vitale, M. Ballardini, V. Borromeo, C. Luzzago, L. Chiavacci, and M. L. Mandola, 2016: Q fever seroprevalence and risk factors in sheep and goats in northwest Italy. *Prev. Vet. Med.* 130, 10–17.

Rockhill, B., B. Newman, and C. Weinberg, 1998: Use and misuse of population attributable fractions. *Am. J. Public Health* 88, 15–19.

Rooney, A. L., G. Limon, H. Vides, A. Cortez, and J. Guitian, 2014: Sarcocystis spp. in llamas (Lama glama) in Southern Bolivia: a cross sectional study of the prevalence, risk factors and loss in income caused by carcass downgrades. *Prev. Vet. Med.* 116, 296–304.

Rose, G., 2001: Sick individuals and sick populations. 1985. *Bull. World Health Organ.* 79, 990–996.

Rossow, H., J. Ollgren, P. Klemets, I. Pietarinen, J. Saikku, E. Pekkanen, S. Nikkari, H. Syrjala, M. Kuusi, and J. P. Nuorti, 2014: Risk factors for pneumonic and ulceroglandular tularaemia in Finland: a population-based case-control study. *Epidemiol. Infect.* 142, 2207–2216.

Rothman, K. J., 2012: Epidemiology: An Introduction. OUP, USA.

Rothman, K. J., and S. Greenland, 1998: Modern Epidemiology, p. xiii, 737 p. Lippincott-Raven, Philadelphia.

Rothman, K. J., and S. Greenland, 2005: Causation and causal inference in epidemiology. *Am. J. Public Health* 95(Suppl 1), S144–S150.

Rothman, K. J., S. Greenland, and A. M. Walker, 1980: Concepts of Interaction. *Am. J. Epidemiol.* 112, 467–470.

Rotnitzky, A., and J. Robins, 1997: Analysis of semi-parametric regression models with non-ignorable non-response. *Stat. Med.* 16, 81–102.

Royston, P., and W. Sauerbrei, 2008: Multivariable Model-Building: A Pragmatic Approach to Regression Anaylsis Based on Fractional Polynomials for Modelling Continuous Variables. John Wiley & Sons, Hoboken, NJ.

Royston, P., G. Ambler, and W. Sauerbrei, 1999: The use of fractional polynomials to model continuous risk variables in epidemiology. *Int. J. Epidemiol.* 28, 964–974.

Royston, P., D. G. Altman, and W. Sauerbrei, 2006: Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat. Med.* 25, 127–141.

Rubin, D. B., 1976: Inference and missing data. *Biometrika* 63, 581–592.

Rubin, D. B., 2008: Introduction. In: Multiple Imputation for Nonresponse in Surveys, pp. 1–26. John Wiley & Sons Inc, Hoboken, NJ.

Sackett, D. L., 1979: Bias in analytic research. *J. Chronic Dis.* 32, 51–63.

Safer, D. J., 2002: Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. *J. Nerv. Ment. Dis.* 190, 583–592.

Saracci, R., 1980: Interaction and Synergism. *Am. J. Epidemiol.* 112, 465–466.

Sargeant, J. M., M. O'Connor A, I. A. Gardner, J. S. Dickson, M. E. Torrence, I. R. Dohoo, S. L. Lefebvre, P. S. Morley, A. Ramirez, and K. Snedeker, 2010: The REFLECT statement: reporting guidelines for randomized controlled trials in livestock and food safety: explanation and elaboration. *Zoonoses Public Health* 57, 105–136.

Sauer, B. C., M. A. Brookhart, J. Roy, and T. VanderWeele, 2013: A review of covariate selection for non-experimental comparative effectiveness research. *Pharmacoepidemiol. Drug Saf.* 22, 1139–1145.

Scantlebury, C. E., D. C. Archer, C. J. Proudman, and G. L. Pinchbeck, 2011: Recurrent colic in the horse: incidence and risk factors for recurrence in the general practice population. *Equine Vet. J. Suppl.* 43(suppl 39), 81–88.

Schafer, J. L., 1997: Analysis of Incomplete Multivariate Data. CRC Press, Boca Raton, FL.

Schemann, K., M. R. Taylor, J. A. L. M. L. Toribio, and N. K. Dhand, 2011: Horse owners' biosecurity practices following the first equine influenza outbreak in Australia. *Prev. Vet. Med.* 102, 304–314.

Scherer, R. W., P. Langenberg, and E. von Elm, 2007: Full publication of results initially presented in abstracts. *Cochrane Database Syst. Rev.* 18, MR000005.

Schisterman, E. F., S. R. Cole, and R. W. Platt, 2009: Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 20, 488–495.

Schukken, Y. H., Y. T. Grohn, B. McDermott, and J. J. McDermott, 2003: Analysis of correlated discrete observations: background, examples and solutions. *Prev. Vet. Med.* 59, 223–240.

Schulz, K. F., and D. A. Grimes, 2002: Case-control studies: research in reverse. *Lancet* 359, 431–434.

Schulz, K. F., and D. A. Grimes, 2005: Sample size calculations in randomised trials: mandatory and mystical. *Lancet* 365, 1348–1353.

Schulz, K. F. , D. G. Altman, D. Moher, and C. Group, 2010: CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *J. Clin. Epidemiol.* 63, 834–840.

Slattery, M. L., S. L. Edwards, B. J. Caan, R. A. Kerber, and J. D. Potter, 1995: Response rates among control subjects in case-control studies. *Ann. Epidemiol.* 5, 245–249.

Smith, G. D., and A. Phillips, 1990: Declaring independence: why we should be cautious. *J. Epidemiol. Community Health* 44, 257–258.

Stelfox, H. T., G. Chua, K. O'Rourke, and A. S. Detsky, 1998: Conflict of interest in the debate over calcium-channel antagonists. *N. Engl. J. Med.* 338, 101–106.

Sturmer, T., and H. Brenner, 2002: Flexible matching strategies to increase power and efficiency to detect and estimate gene-environment interactions in case-control studies. *Am. J. Epidemiol.* 155, 593–602.

Szklo, M., and F. J. Nieto, 2000: Epidemiology, Beyond the Basics. Jones and Bartlett, Sandbury (MA).

Thumbi, S. M., B. M. Bronsvoort, E. J. Poole, H. Kiara, P. G. Toye, M. N. Mbole-Kariuki, I. Conradie, A. Jennings, I. G. Handel, J. A. Coetzer, J. C. Steyl, O. Hanotte, and M. E. Woolhouse, 2014: Parasite co-infections and their impact on survival of indigenous cattle. *PLoS ONE* 9, e76324.

Tobias, T. J., D. Klinkenberg, A. Bouma, J. van den Broek, A. J. Daemen, J. A. Wagenaar, and J. A. Stegeman, 2014: A cohort study on Actinobacillus pleuropneumoniae colonisation in suckling piglets. *Prev. Vet. Med.* 114, 223–230.

Trehy, M. R., A. J. German, P. Silvestrini, G. Serrano, and D. J. Batchelor, 2014: Hypercobalaminaemia is associated with hepatic and neoplastic disease in cats: a cross sectional study. *BMC Vet. Res.* 10, 175.

Uter, W., and A. Pfahlberg, 2001: The application of methods to quantify attributable risk in medical practice. *Stat. Methods Med. Res.* 10, 231–237.

Vandenbroucke, J. P., 1991: Prospective or retrospective: what's in a name? *BMJ* 302, 249–250.

Vandenbroucke, J. P., E. von Elm, D. G. Altman, P. C. Gotzsche, C. D. Mulrow, S. J. Pocock, C. Poole, J. J. Schlesselman, M. Egger, and S. Initiative, 2007: Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Epidemiology* 18, 805–835.

Vandenbroucke, J. P., E. von Elm, D. G. Altman, P. C. Gotzsche, C. D. Mulrow, S. J. Pocock, C. Poole, J. J. Schlesselman, M. Egger, and S. Initiative, 2014: Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int. J. Surg.* 12, 1500–1524.

Wakefield, M. A., F. J. Chaloupka, N. J. Kaufman, C. T. Orleans, D. C. Barker, and E. E. Ruel, 2000: Effect of restrictions on smoking at home, at school, and in public places on teenage smoking: cross sectional study. *BMJ* 321, 333–337.

Waldner, C. L., 2008: Western Canada study of animal health effects associated with exposure to emissions from oil and natural gas field facilities. Study design and data collection I. Herd performance records and management. *Arch. Environ. Occup. Health* 63, 167–184.

Waldner, C. L., and L. B. Rosengren, 2009: Factors associated with serum immunoglobulin levels in beef calves from Alberta and Saskatchewan and association between passive transfer and health outcomes. *Can. Vet. J.* 50, 275–281.

Ward, M. P., L. T. Glickman, and L. E. Guptill, 2002: Prevalence of and risk factors for leptospirosis among dogs in the United States and Canada: 677 cases (1970-1998). *J. Am. Vet. Med. Assoc.* 220, 53–58.

Watters, M. E., K. M. Meijer, H. W. Barkema, K. E. Leslie, M. A. von Keyserlingk, and T. J. Devries, 2013: Associations of herd- and cow-level factors, cow lying behavior, and risk of elevated somatic cell count in free-stall housed lactating dairy cows. *Prev. Vet. Med.* 111, 245–255.

Weng, H. Y., P. H. Kass, L. A. Hart, and B. B. Chomel, 2006: Risk factors for unsuccessful dog ownership: an epidemiologic study in Taiwan. *Prev. Vet. Med.* 77, 82–95.

Wieland, S., and K. Dickersin, 2005: Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *J. Clin. Epidemiol.* 58, 560–567.

Willeberg, P., L. R. Nielsen, and M. Salman, 2012: Designing and evaluating risk-based surveillance systems: Potential unwarranted effects of applying adjusted risk estimates. *Prev. Vet. Med.* 105, 185–194.

Williams, K. J., M. P. Ward, O. P. Dhungyel, and E. J. Hall, 2015: Risk factors for Escherichia coli O157 shedding and super-shedding by dairy heifers at pasture. *Epidemiol. Infect.* 143, 1004–1015.

Zhao, L. P., and L. N. Kolonel, 1992: Efficiency loss from categorizing quantitative exposures into qualitative exposures in case-control studies. *Am. J. Epidemiol.* 136, 464–474.

