ORIGINAL ARTICLE

The REFLECT Statement: Reporting Guidelines for Randomized Controlled Trials in Livestock and Food Safety: Explanation and Elaboration[†]

J. M. Sargeant[1], A. M. O'Connor[2], I. A. Gardner[3], J. S. Dickson[4], M. E. Torrence[5] and consensus meeting participants*: I. R. Dohoo[6], S. L. Lefebvre[7], P. S. Morley[8], A. Ramirez[2] and K. Snedeker[1]

[1] Centre for Public Health and Zoonoses, and Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, ON, Canada
[2] Veterinary Diagnostic and Production Animal Medicine, Veterinary Medicine Research Institute Building 4, Iowa State University, Ames, IA, USA
[3] Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, CA, USA
[4] Iowa State University, Department of Animal Science, 215F Meat Laboratory, Ames, IA, USA
[5] USDA-ARS, GWCC-4-2194, Beltsville, MD, USA
[6] Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, PE, Canada
[7] American Veterinary Medical Association, Schaumburg, IL, USA
[8] Department of Clinical Sciences, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO, USA

Impacts

- Complete and accurate reporting of randomized controlled trials is necessary to allow the reader of the trial to evaluate internal and external validity.
- The REFLECT statement provides a checklist of items to include when reporting randomized controlled trials conducted in livestock and food safety.
- This explanation and elaboration document provides details for trial authors and reviewers using the REFLECT statement checklist.

Abstract

Concerns about the completeness and accuracy of reporting of randomized clinical trials (RCTs) and the impact of poor reporting on decision making have been documented in the medical field over the past several decades. Experience from RCTs in human medicine would suggest that failure to report critical trial features can be associated with biased estimated effect measures, and there is evidence to suggest that similar biases occur in RCTs conducted in livestock populations. In response to these concerns, standardized guidelines for reporting RCTs were developed and implemented in human medicine. The Consolidated Standards of Reporting Trials (CONSORT) statement was first published in 1996, with a revised edition published in 2001. The CONSORT statement consists of a 22-item checklist for reporting a RCT and a flow diagram to follow the number of participants at each stage of a trial. An explanation and elaboration document not only defines and discusses the importance of each of the items, but also provides examples of how this information could be supplied in a publication. Differences between human and livestock populations necessitate modifications to the CONSORT statement to maximize its usefulness for RCTs involving livestock. These have been addressed in an extension of the CONSORT statement titled the REFLECT statement: Methods and processes of creating reporting guidelines for randomized control trials for livestock and food safety. The modifications made for livestock trials specifically addressed the common use of group housing and group allocation to intervention in livestock studies; the use of deliberate challenge models in some trials

and the common use of non-clinical outcomes, such as contamination with a foodborne pathogen. In addition, the REFLECT statement for RCTs in livestock populations proposed specific terms or further clarified terms as they pertained to livestock studies.

The randomized clinical trial is a very beautiful technique, of wide applicability, but as with everything else, there are snags. When humans have to make observations, there is always the possibility of bias (Cochrane, 1972).

The randomized controlled trial (RCT) is the gold standard for evaluating the efficacy of therapeutic and preventive interventions. In livestock populations, RCTs can be used to evaluate the efficacy of interventions related to animal health and productivity, as well as food-safety outcomes. However, trials that do not employ sound methodologies are associated with biased-effect estimates (Schulz et al., 1995; Moher et al., 1998; Juni et al., 2001). Biased trial results have the potential to mislead decision making by clinicians, researchers and policy makers, which ultimately impacts livestock producers and the general public. The reader of a published clinical trial cannot know the exact methods used to conduct the trial, as the only information available to the reader is that provided in the publication. Therefore, it is essential that authors of clinical trials provide complete and accurate details of the methods used in the trials in the publication.

Incomplete and inaccurate reporting in published livestock intervention trials

The basic criteria essential to the validity of RCTs have been reviewed in the veterinary literature (Ribble, 1990; Lund et al., 1994; Dohoo, 2004). However, despite the availability of these criteria, the quality of reporting of intervention trials remains poor. An assessment of the quality of RCTs published in one journal revealed that although some of the trials provided information on methodological features, many others failed to do so (Elbers and Schukken, 1995). These trials lacked information related to the method of treatment allocation, the grouping of animals relative to treatment allocation, the use or non-use of blinding and the method of statistical analysis (Elbers and Schukken, 1995). Further, several systematic reviews in pre-harvest food safety (Denagamage et al., 2007; O'Connor et al., 2008) and animal health (O'Connor et al., 2006; Wellman and O'Connor, 2007; Burns and O'Connor, 2008) have noted a lack of reporting of group allocation methods; blinding and details related to intervention protocols, outcome assessments

and statistical analysis methods in some published clinical trials. This lack of consistency in reporting makes it almost impossible to summarize sufficient data appropriately, thereby affecting the ability to arrive at an overall conclusion on a particular intervention or outcome. For example, in 100 randomly selected trials on animal health or production outcomes, only 67% reported random allocation to intervention group, 35% clearly described the number of animals housed together in a group, 4% reported the use of double blinding where blinding was feasible and 62% reported the number of study units lost to follow-up during the trial (Sargeant et al., 2009a). In an evaluation of 100 pre-harvest food-safety trials, randomization, double blinding and the number of subjects lost to follow-up were reported in 46%, 0% and 43% of trials respectively, and the number of animals housed together was stated in 52% of the trials (Sargeant et al., 2009b).

Experience gained from RCTs in human medicine would suggest that failure to report critical trial features can be associated with biased estimates of effect measures, and there is evidence to suggest that similar biases occur in RCTs conducted in livestock populations. A systematic review of trials evaluating the efficacy of vaccination for the treatment of pink-eye in cattle found that trials not reporting random allocation to intervention group and blinding were more likely to conclude that the vaccine was efficacious than trials where these features were reported (Burns and O'Connor, 2008). Similarly, evaluations of 100 randomly selected trials with animal health or production outcomes and 100 randomly selected trials with food-safety outcomes revealed significant associations between the proportion of positive treatment effects within trials and failure to report trial features, such as random allocation to intervention group, exclusion criteria for study subjects, details of the intervention protocol, animal signalment and details of the measurement of all outcomes (Sargeant et al., 2009a,b).

Improving the reporting of RCTs in the medical literature: The CONSORT statement

Concerns about the completeness and accuracy of reporting of RCTs have been documented in the medical field over the past several decades (DerSimonian et al., 1982; Pocock et al., 1987; Gotzsche, 1989; Schulz et al., 1994;

Sonis and Joines, 1994; Ah-See and Molony, 1998). In response to these concerns, standardized guidelines for reporting RCTs were developed and have been implemented. The Consolidated Standards of Reporting Trials (CONSORT) statement was first published in 1996 (Begg et al., 1996). A revised version was simultaneously published by four leading medical journals in 2001 (Moher et al., 2001b,c,d,e). The CONSORT statement consists of a 22-item checklist for reporting an RCT and a flow diagram to follow the number of participants at each stage of a trial. The items for the checklist were selected because there was empirical evidence in the literature indicating the potential for biased estimates of treatment effects when these items were not reported, or because the information was deemed essential to evaluate the reliability or relevance of the findings (Moher et al., 1998). An explanation and elaboration document not only defines and discusses the importance of each of the items, but also provides examples of how this information could be supplied in a publication (Altman et al., 2001). The CONSORT statement document is currently endorsed by several hundred journals (http://www.consort-statement.org), including two veterinary journals: the *Equine Veterinary Journal* and *The Veterinary Journal* (Higgins, 1997). Evaluations of RCTs since implementation of the CONSORT statement suggest that the statement has improved the quality of reporting of RCTs (Moher et al., 2001a; Plint et al., 2006; Kane et al., 2007). Extensions of the CONSORT statement have been developed for cluster trials (Campbell et al., 2004a, 2005, 2006), harms (Ioannidis et al., 2004), herbals interventions (Gagnier et al., 2005, 2006a,b,c), non-pharmacological interventions (Boutron et al., 2008) and abstracts (Hopewell et al., 2008).

Modifications to the CONSORT statement for use in trials involving livestock species

Differences between human and livestock populations necessitate modifications to the CONSORT statement to maximize its usefulness for RCTs involving livestock. These have been addressed in an extension of the CONSORT statement titled the 'REFLECT' statement – The REFLECT statement: Methods and processes of creating reporting guidelines for randomized control trials for livestock and food safety (O'Connor et al., 2010a,b,c,d,e). The modifications to the CONSORT checklist recommended for livestock populations in the REFLECT statement for livestock and food-safety intervention studies are presented in Table 1. Although many of the checklist items from the CONSORT statement remain unchanged, the modifications made for documentation of livestock trials (O'Connor et al., 2010a,b,c,d,e) specifically addressed

the common use of group housing and group allocation to intervention in livestock studies, the use of deliberate challenge models in some trials and the common use of non-clinical outcomes, such as contamination with a food-borne pathogen. In addition, the REFLECT statement for RCTs in livestock populations proposed specific terms or further clarified terms as they pertained to livestock populations. The term 'participant' in the original CONSORT statement was limited to refer only to animals' owners/managers, who consent to participate in the trial. The term 'study unit' was preferred and recommended in the REFLECT statement for the units within the study. This term was used instead of 'animal unit', as it is common that a part of an animal, such as a hoof, teat or eye, be allocated to treatment. Study units may further be classified as allocation units and outcome units. For example, a study may allocate udder halves to receive the treatment; therefore, the allocation unit is the udder half. However, the outcome may be measured on the individual teat (i.e. the outcome unit).

The objective of this explanation and elaboration document is to define each item modified from the CONSORT checklist for the REFLECT statement for livestock and food safety, to provide a rationale for its inclusion and to provide illustrative examples of how the item might be reported for each REFLECT item. The examples are derived from previously published studies in the animal health/production and pre-harvest food-safety literature.

Definitions

Challenge trial
A study design where the investigator controls allocation to intervention and disease occurrence. In therapeutic challenge trials, the investigator uses a model to induce disease, and then allocates the study units to receive the therapeutic intervention. The outcome of interest is often clinical improvement. In therapeutic challenge trials with health and production outcomes, the condition of interest is commonly exposure to an infectious pathogen or a metabolic disease, such as fatty liver in dairy cattle.

In preventive challenge trials, the investigator allocates the study units to receive the preventive intervention, and then uses a disease model to induce disease. The outcome of interest is often prevention of clinical signs. For preventive challenge studies with food-safety outcomes, the study often ensures exposure to the pathogen of interest. Although challenge trials do not always involve an infectious-disease outcome, this is a common model in livestock populations and therefore, throughout the text, most references to challenge trials are limited to infectious-agent models.

Table 1. Checklist of Items for the REFLECT statement: reporting guidelines for randomized control trials in livestock and food safety

Paper section and topic	Item	Descriptor of REFLECT statement item	Reported on page no.
Title & Abstract	1	How study units were allocated to interventions (e.g. 'random allocation', 'randomized' or 'randomly assigned'). **Clearly state whether the outcome was the result of natural exposure or was the result of a deliberate agent challenge**	
Introduction Background	2	Scientific background and explanation of rationale	
Methods Participants	3	Eligibility criteria **for owner/managers and study units at each level of the organizational structure,** and the settings and locations where the data were collected	
Interventions	4	Precise details of the interventions intended for each group, **the level at which the intervention was allocated** and how and when interventions were actually administered	
	4b	**Precise details of the agent and the challenge model, if a challenge study design was used**	
Objectives	5	Specific objectives and hypotheses. **Clearly state primary and secondary objectives (if applicable)**	
Outcomes	6	Clearly defined primary and secondary outcome measures and the levels at which they were measured and, when applicable, any methods used to enhance the quality of measurements (e.g. multiple observations and training of assessors)	
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules. **Sample size considerations should include sample size determinations at each level of the organizational structure and the assumptions used to account for any non-independence among groups or individuals within a group**	
Randomization – Sequence generation	8	Method used to generate the random allocation sequence **at the relevant level of the organizational structure**, including details of any restrictions (e.g. blocking and stratification)	
Randomization – Allocation concealment	9	Method used to implement the random allocation sequence **at the relevant level of the organizational structure**, (e.g. numbered containers ~~or central telephone~~), clarifying whether the sequence was concealed until interventions were assigned	
Randomization – Implementation	10	Who generated the allocation sequence, who enrolled **study units** and who assigned **study units** to their groups **at the relevant level of the organizational structure**	
Blinding (masking)	11	Whether or not ~~participants~~ those administering the interventions, **caregivers** and those assessing the outcomes were blinded to group assignment. If carried out, how the success of blinding was evaluated**. Provide justification for not using blinding if it was not used**	
Statistical methods	12	Statistical methods used to compare groups for all outcome(s). Clearly state the level of statistical analysis **and methods used to account for the organizational structure, where applicable;** methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Study flow	13	Flow **of study units** through each stage **for each level of the organization structure of the study** (a diagram is strongly recommended). Specifically, for each group, report the numbers of **study units** randomly assigned, receiving intended treatment, completing the study protocol and analysed for the primary outcome. Describe protocol deviations from study as planned, together with reasons	
Recruitment	14	Dates defining the periods of recruitment and follow-up.	

Table 1. (*Continued*)

Paper section and topic	Item	Descriptor of REFLECT statement item	Reported on page no.
Baseline data	15	Baseline demographic and clinical characteristics of each group, **explicitly providing information for each relevant level of the organizational structure. Data should be reported in such a way that secondary analysis, such as risk assessment, is possible**	
Numbers analysed	16	Number **of study units** (denominator) in each group included in each analysis and whether the analysis was by 'intention to treat'. State the results in absolute numbers when feasible (e.g. 10/20, not 50%).	
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group, **accounting for each relevant level of the organizational structure**, and the estimated effect size and its precision (e.g. 95% confidence interval)	
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory	
Adverse events	19	All important adverse events or side effects in each intervention group	
Discussion			
Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes. **Where relevant, a discussion of herd immunity should be included. If applicable, a discussion of the relevance of the disease challenge should be included**	
Generalizability	21	Generalizability (external validity) of the trial findings	
Overall evidence	22	General interpretation of the results in the context of current evidence	

Text in bold are modifications from the original CONSORT description.

Study unit

The term 'study unit' refers to the units within the study; synonyms may be the 'unit of concern' or 'experimental unit'. Examples of study units may be a hoof, teat, eye, animal, pen or barn.

Allocation unit

This term refers to the study unit that is randomly allocated to receive the intervention. The allocation unit can occur at only one level of the organizational structure.

For example, in a swine study evaluating the impact of a water-based vaccine on weight gain, barns may be randomly allocated to receive the water-based vaccine or a placebo; therefore, the allocation unit is the barn. In a challenge study evaluating the impact of a chilling process intervention on the prevalence of *Campylobacter* on poultry carcasses, carcass halves may be randomly allocated to receive either processing method A or B; therefore, the allocation unit is the carcass half.

Outcome unit

This term refers to the unit at which outcomes are measured. Common outcomes in livestock production are weight gain, disease occurrence or the presence or absence of an infectious disease agent. The outcome unit can occur at only one level of the organizational structure, and may be at the same level of the organizational structure as the allocation unit, or at a lower level. For example, in a swine study evaluating the impact of a water-based vaccine on weight gain, barns may be randomized to receive the intervention; therefore, the unit of allocation is the barn. If weight gain was measured by weighing all animals in the pen on a group scale at the end of the study period (i.e. individual weights are not available), then the outcome unit is the pen. Alternatively, if weight gain is measured by weighing each animal individually, then the outcome unit is the animal, i.e. there are multiple outcome units within the allocation unit. However, in a challenge study evaluating the impact of a chilling process on the prevalence of *Campylobacter* on poultry carcasses, carcass halves may be randomly allocated to receive the intervention. If the presence or absence of *Campylobacter* is also measured on carcass halves, then the outcome unit is the carcass half, which is also the allocation unit.

Primary outcome

The primary outcome refers to an outcome variable of interest, the expected value of which is used to determine

the study sample size. If researchers have more than one outcome of interest, the sample size will be determined by the outcome that needs the highest sample size, and this will be the primary outcome.

Secondary outcome(s)

This refers to another outcome measure that is potentially equally important but not used to determine the sample size. There may be more than one secondary outcome.

Level of organizational structure: The level of organizational structure refers to the manner in which the allocation and outcome units are organized within a production system. The organizational structure may not always be hierarchical (i.e. not always nested).

Examples of organizational structure
In a swine study evaluating the impact of a vaccine on piglet mortality, the animals may be at the bottom of an organizational structure that could include: (1) the production company, (2) the site within the production company, (3) the barn within the site, (4) the pen/room within the barn, (5) the sow within the room and (6) the piglet within the sow's litter. In this example, a hierarchy, or nested structure, is apparent.

In a feedlot-based cattle study evaluating the impact of metaphylaxis with an injectable antibiotic on the occurrence of respiratory disease in cattle, the cattle may be at the bottom of an organizational structure that could include: (1) the originating farm or order buyer, (2) the receiving feedlot, (3) the truckload and (4) the pen. In this situation, the nested hierarchy apparent in the piglet example (see above) does not exist, as different order buyers may have multiple truckloads, which are mixed in different pens.

REFLECT Checklist Items

In this section, square brackets ([]) indicate that explanatory information has been inserted into the quoted text by the REFLECT statement authors to clarify the quoted text. Citations originally included in the quoted text have been removed to avoid confusion.

Title and abstract

Item 1
How study units were allocated to interventions (e.g. 'random allocation', 'randomized' or 'randomly assigned'). Clearly state whether the outcome was the result of natural exposure or the result of a deliberate agent challenge.

Examples

A randomized herd-level field study of dietary interactions with monensin on milk fat percentage in dairy cows (Dubuc et al., 2009).

Efficacy of a novel trivalent inactivated vaccine against the shedding of *Salmonella* in a chicken challenge model (Deguchi et al., 2009).

Explanation

Citation databases frequently search for citations on the basis of abstract and title. The inclusion of terms that include the word 'random', such as 'random allocation', 'randomized', 'randomization' or 'randomly assigned' in the title and/or abstract will allow easy identification of this study design for people conducting electronic data searches to identify evidence for the efficacy of interventions, and for those conducting systematic reviews.

Further, there are important differences with respect to the external validity of studies using models of disease, as occurs in challenge trials, versus natural development of the disease, as occurs in field trials, particularly for infectious diseases. Challenge trials are often conducted under controlled experimental conditions, with a single pathogen in a restricted population. Consequently, the external validity of the challenge study may not compare favourably with the same trial conducted under commercial conditions using a natural disease exposure. Therefore, the identification of a trial as having used natural or deliberate exposure allows for the rapid differentiation of these studies. We strongly encourage the use of the term 'field trial' or 'clinical trial' to describe studies associated with natural development of the disease and the term 'challenge study' or 'challenge trial' or 'challenge model' for studies that use induced models of disease.

Introduction

Item 2
Scientific background and explanation of rationale.

Example

The success of commercial dairies depends on a reliable supply of healthy replacement heifer calves with good genetic potential for milk production. Several management practices have been recommended to producers for reducing the frequency of calf morbidity and mortality on dairy farms. One area commonly emphasized is the calving pen. The

management of calving pens influences the degree of early calf exposure to infectious environmental pathogens (Pithua et al., 2009).

Explanation

The introduction should provide sufficient contextual background, as it relates to the study topic, to provide the reader with a basic understanding of the underlying science upon which the study was based. This should include a description of the nature, scope and extent or magnitude of the problem under study; the pathophysiological basis for active components in the proposed treatment or the justification for considering a new treatment regimen when there is an existing treatment, as well as any other factors known to influence the outcome and interpretation of data for the study topic.

Authors should indicate whether the intervention is directed at a single component or multiple components associated with the aetiology of the naturally occurring disease. For instance, challenge trials or field trials may test the efficacy of an intervention against specific bacteria, whereas natural development of the disease may be associated with multiple organisms. Providing this information in the introduction provides the reader with the context necessary for the interpretation of the study results.

The introduction section should also provide a rationale justifying the need for the research. This may include an identification of knowledge gaps, as well as an indication as to how the current study will enhance our knowledge in the topic area. The authors should provide an overview of the current state of knowledge, based on other published studies. If available, the authors should reference any systematic reviews completed for the same or related interventions. The CONSORT statement (Moher et al., 2001d) suggests that for some human disease processes, a formal review of the published literature may be the preferred course of action over carrying out another (unnecessary) primary study. In livestock species, there is a paucity of primary studies for many interventions, and systematic reviews are not yet commonly used (Sargeant et al., 2006).

Many veterinary and food-safety journals prefer that the specific objectives be included in the final paragraph of the introduction section. In the CONSORT statement, the objectives were described in Item 5 in the methods and materials, and the REFLECT statement left the item relating to the study objectives as Item 5, although we recognize that the introduction often will be an appropriate place for this information.

Methods

Item 3
Eligibility criteria for owners/managers and study units at each level of the organizational structure, and the settings and locations where the data were collected.

Examples of eligibility criteria

Study farms were initially identified through private veterinary practices (PVP), which had submitted any kind of cattle samples for diagnosis to the Veterinary Laboratories Agency's regional laboratories (VLA RL) during the previous 12 months as previously described. The cattle farms within each PVP, who submitted the largest number of samples in the previous year, were included and further suggestions of potentially interested farmers from the PVP were also accepted. Neighbouring farms were excluded. A total of 411 farms distributed throughout England and Wales were contacted by phone to assess willingness to participate in the study and eligibility of the herd by questionnaire. Farms were eligible, if they retained more than 60 cattle including 20 young stock, had a bovine tuberculosis-negative status, and the premises were not shared with any public access enterprises such as open farms, Bed & Breakfast or farm-shops including selling unpasteurized milk (Ellis-Iversen et al., 2008).

Animals that arrived at the feedlot between October 16, 1994, and December 13, 1994, were candidates for the trial. In this study, the case definition for UF [undifferentiated fever] was an elevated rectal temperature (>40.5°C) and a lack of abnormal clinical signs referable to organ systems other than the respiratory system within 3 wk after arrival at the feedlot. Exclusion criteria were moribund animals and animals with a previous treatment history for any disease (Jim et al., 1999).

Explanation

All trials address an issue relevant to a population of interest, i.e. the target population; however, for logistic reasons, trials use eligibility criteria to define a study population. Study unit selection on the basis of eligibility criteria may lead to meaningful differences between the target population and the study population; therefore, these eligibility criteria must be stated explicitly to enable the reader to assess differences between the study and target populations, and ultimately to assess external validity of findings. It is not necessary to describe both eligibility

and exclusion criteria, as study units that do not fit the eligibility criteria are excluded.

In the human medical field, this item generally relates to eligibility criteria for participants and restriction of the trial setting to one or more medical centres (Altman et nbsp;al., 2001). In livestock trials, the concept of 'participant' refers to the owner or manager of the animals who consents to participate in the trial. Thus, it is important to report eligibility criteria of the owner/manager and also eligibility criteria for the study units. Livestock studies frequently need to consider multiple levels of organizational structure when the study units are enrolled. For example, for evaluation of the efficacy of swine vaccines, the following are usually enrolled: owners of the facilities, barns within the facilities, pens within the barns and finally pigs within the pens. Decisions made about eligibility criteria at each organizational level may influence differences between the study population and the target population, and should be reported.

Frequently, the only determinant of eligibility for a facility may be a personal relationship with the researcher or a veterinary practice and willingness to co-operate by the owners/managers, or the proximity to the researchers' laboratories. If such convenience sampling is used, this should be stated. In other situations, facilities may be selected randomly from a sampling frame, such as a premises identification database or livestock commodity organization or program list. In some instances, farms may be selected on the basis of the presence or frequency of occurrence of the disease of interest.

At the study-unit level, eligibility criteria commonly include age or production stage, sex, co-morbidities or previous treatments. For example, it is common for livestock-production trials to exclude study units with a prior history of the disease of interest, i.e. excluding animals with an existing antibody titre to a specific pathogen in trials that are evaluating the efficacy of a vaccine to prevent illness caused by that pathogen. In challenge studies, it is common that only animals not colonized by the pathogen of interest are eligible for the study, i.e. swine colonized with *Salmonella* may be excluded from a study planning to use an artificial challenge with *Salmonella*.

Examples of setting and location information

The experiment was carried out in a mountainous area (1,000 m above sea level) in the northwest of Spain (6°53′W, 43°21′N; Sierra de San Isidro, Illano, Asturias), where shrubby heather-gorse vegetation is dominant. Four plots of 5000 m² each were established, in which the vegetation had been improved in 2001 by soil ploughed and dressing and sowing perennial ryegrass (*Lolium perenne* L.) and white clover (*Trifolium repens* L.), and removing any heather that was present. Annual rainfall in the experimental year (2004) was 1,589 mm. During the grazing season, mean rainfall ranged from 36 to 111 mm/mo. Mean average temperatures were 17.3°-C in June and 10.6°-C in May (Osoro et al., 2007).

Broiler chicks were hatched from commercially obtained eggs and grown to market age (56 to 63 d) on pine shavings in floor pens (5 × 8 m) in a controlled environment-type house....All broilers were processed in the pilot plant processing facility at the Russell Research Center (Northcutt et al., 2006).

The setting and location may affect the external validity of the study. For some diseases, it may be relevant to report the geographical location(s) where the trial was conducted, as the frequency of many livestock diseases and the response to interventions varies geographically as a result of differences in climate and management systems. The time of year when the study was conducted may also be relevant to disease frequency. When reporting time of year, the month(s) and year should be included, and the reader should be allowed to infer the season.

At the farm level, issues related to setting that could influence the external validity of the study should be described. Authors should describe the group sizes for all relevant levels of the organizational structure, i.e. the capacity of the facility and the number, size and capacity of barns/pens/cages, etc., used to house study units. Feed and other pertinent management details and the presence or absence of the disease of interest, or other endemic diseases, should also be described. The nature of the management of the facility should also be reported. As an example, there may be differences in facility management between a commercial operation of a large company; an independent, privately owned facility and a facility operated by a university or government research organization.

Item 4a

Precise details of the interventions intended for each group, at the level at which the intervention was allocated, and how and when interventions were actually administered.

Example

Treatment was assigned at the heifer level within herd. Each heifer was randomly assigned to a treatment using a random number generator function (R Development Core Team, 2006), and the farmers were blinded to the treatment. Before treatment, each teat-end was scrubbed with a cotton wool

pledget moistened in 70% methanol and a gland secretion sample was collected aseptically ($n = 4{,}268$ glands). No secretion was discarded before collection because there was only a small total volume of secretion present in most glands. If no secretion could be collected from a gland, it was recorded as a missing sample ($n = 99$ glands). Following sampling, all 4 glands within a heifer were infused with 2.6 g of bismuth subnitrate following teat-end scrubbing ($n = 268$ heifers; Teat Seal, Pfizer Animal Health NZ Ltd., Auckland, New Zealand), or a heifer was administered with 5 g of tylosin base i.m. for 3 d at 24-h intervals ($n = 268$ heifers; Tylan 200, Elanco Animal Health, Manukau City, New Zealand), or all 4 glands were infused with the teat sealant and the heifer was administered 5 g of tylosin base i.m. for 3 d at 24-h intervals ($n = 266$ heifers), or they were left as an untreated control ($n = 265$ heifers). The tip of the teat sealant cannula was inserted approximately 3 mm into the teat canal for infusion. Following sampling or infusion, 0.5% effective iodine was applied by manual spraying to all teat ends. Technicians administered the first treatment of tylosin and then left labeled doses of tylosin for the remaining 2 treatments for farm staff to administer (Parker et al., 2008).

Explanation

The description of the intervention(s), including the control intervention, should be provided in sufficient detail to allow the reader to replicate the intervention. Phrases such as 'applied per labelled instructions', 'as per manufacturers' instructions', 'standard industry practices' or 'routine treatment' do not constitute an adequate description that can be replicated. Differences in management or handling among intervention groups should be included in the description of the interventions.

The unit of allocation for the intervention(s) should be clearly stated, and this unit must correspond with the unit of randomization. Examples of phrases to be used include 'the barn was randomly allocated to receive either treatment A or treatment B', 'the site was randomly allocated to receive either treatment A or treatment B' or 'the teat was randomly allocated to receive either treatment A or treatment B'. These phrases will eliminate confusion often associated with current descriptions. The intent is to state clearly the unit of allocation with adequate detail, so that there is no ambiguity for the reader of the trial report.

For pharmaceutical interventions, the minimum description should include the compound name, the concentration, the dose, the delivery matrix and the route and the frequency of administration.

For biological interventions such as vaccinations, the minimum description should include the organism(s) and whether each one is a modified-live or killed product, substance or probiotic unit; the adjuvant; the concentration per ml (if known); the dose; the delivery matrix and the route and the frequency of administration.

For surgical interventions, the minimum description should include the training level of the person administering the procedure, the number of people administering each procedure, the prior number of times the person had performed the procedure and the post-operative care, including the use of other post-operative treatments such as antibiotics or medications for alleviation of pain. For example, a field trial comparing surgical versus toggle (non-surgical) repair of a left-displaced abomasum repair should include a complete description of the surgical procedure, including post-operative care and how that care differed from the post-operative treatment of cases receiving toggle intervention. For a surgical intervention, it is important to include who performed the procedure, as a procedure performed by farm staff versus a veterinarian may represent different interventions.

For food-processing interventions, the minimum description should include the production process and variables that may affect the outcome of that process. For example, an intervention assessing chlorine concentrations during immersion chilling in a poultry plant should describe the volume of water per carcass, the water refresh rate, the water pH, the water temperature, the water hardness, available chlorine versus total chlorine concentration, the source of the chlorine and the length of time of carcass immersion for each intervention.

It is also preferable to state clearly whether treatments groups are similar, instead of leaving it to the reader to assume that the groups are the same with respect to other factors that could affect the outcome. For example, in a feedlot trial assessing the pen-level prevalence of *Escherichia coli* 0157 in pens that received probiotic A compared with probiotic B at arrival, it is preferable to state clearly that all animals received the same ration or water from the same water supply, if ration or water supply are thought to impact *E. coli* 0157 prevalence.

If the intervention was applied to individual animals, the authors should state whether the animals were individually housed or housed in a group, and if so, the number per housing group. If the intervention was applied at the group level, the authors should clearly state the number of animals per group. The description of housing of the allocation units should correspond with the levels of the organizational structure described in Item 3.

Information about the housing of the allocation units should be described, as this information is essential for assessing the appropriateness of the statistical analysis and the external validity of the study. This information will also further clarify whether the study was a field trial under normal production conditions, a field trial using small numbers of animals per pen (as is common in trials conducted in research herds) or a controlled study under laboratory conditions.

In some challenge trials, non-challenged animals are included to serve as negative controls. When this is a feature of the trial, the number of negative controls and their housing relative to the study units (i.e. within challenged groups, or proximity to, and opportunity for contact with, challenged animals) should be described.

Item 4b

Precise details of the agent and the challenge model, if a challenge study design was used.

Examples

A mixture of three *E. coli* O157:H7 strains resistant to 50 g mL–1 nalidixic acid was used as inoculum for the experiment with sheep. The mixture contained *E. coli* O157:H7 strains E32511 and E318N (human isolates), and H4420nal (bovine isolate). The three strains were cultured individually in tryptic soy broth for 18 to 24 h at 37°C (200 rpm). The optical density (OD640) was measured to ensure approximately equal cell density of all cultures. Aliquots (4 ml) of each strain were pooled with 13 ml of sterile PBS (pH 7.4) in sterile 60-mL polypropylene containers. Subsamples were serially diluted in PBS and enumerated by plating 100 μl aliquots in duplicate onto sorbitol MacConkey agar amended with cefixime (50 mg L^{-1}), potassium tellurite (2.5 mg L^{-1}) and nalidixic acid (50 μg ml^{-1}), denoted CT-SMACnal.... Feed was withdrawn 48 h before inoculation to promote establishment of the inoculated *E. coli* O157:H7 in the gastrointestinal tract. On day 0, each sheep was orally inoculated with 10^{10} CFU of the three-strain mixture of *E. coli* O157:H7 using a 60-mL syringe connected to a polypropylene orogastric tube. The inoculum was followed by two 60 mL aliquots of sterile PBS to rinse the syringe and tubing. Faecal samples were collected from each animal on day 1, to confirm shedding of nalidixic acid-resistant (nalR) *E. coli* O157:H7 (Cook et al., 2005).

The animals were divided into 2 groups of 12 cows each (6 pairs per group) that went through the protocol 4 wk apart. The duration of the experimental period was 17 d. From d 0 to 6, cows were fed a standard diet based on a forage mix with 50% alfalfa silage (42% NDF and 21% CP on a DM basis) and 50% corn silage (38% NDF and 9% CP on a DM basis) fed *ad libitum*, and offered twice a day, at 0730 and 1500 h. Nutrient composition of forage was determined on a 6500 NIR spectrophotometer (Foss in North America, Eden Prairie, MN) using equations of the NIRS Consortium. Vitamins and minerals were fed to meet requirements, mixed with 1.4 kg of corn-based concentrate. Vitamins represented 1.0% of the DM of the concentrate (3,304 IU/g of DM of vitamin A, 1,101 IU/g of DM of vitamin D, and 55 IU/g of DM of vitamin E) and minerals represented 0.6% of the DM of the concentrate (0.55% Mn, 0.55% Zn, 0.35% Fe, 0.14% Cu, 0.008% I, 0.006% Se, and 0.002% Co). On d 7, cows were restricted to 30% of the energy required for pregnancy and maintenance by restricting the intake of a forage mix, based on equal proportions of alfalfa silage, corn silage, and wheat straw that was offered once a day in the morning in addition to the 1.4 kg of concentrate previously described. Wheat straw analysis indicated a CP content of 3.5% and 77% NDF (Cooke et al., 2007).

Explanation

The precise details of the challenge model used in the study are critically important for assessing the external validity (Item 21). Challenge trials represent an enormously broad spectrum of conditions. Often, challenge trials involve exposure to infectious agents. These models of disease may not always be associated with clinical disease; for example, challenge models of foodborne pathogens rarely induce clinical disease. Other models may not have an infectious component, such as lameness models or models of metabolic disease, such as fatty liver in dairy cattle. The onus is on the authors to provide sufficient details of the model used in the challenge trial to enable the reader to assess its validity as a model for the 'real' condition.

It is not possible to provide guidelines that adequately describe all possible models. However, for an infectious model, it is recommended that the following be included:

1 The timing of challenge relative to intervention, i.e. X h prior to initiation of the intervention (for therapeutic interventions), or X h after the intervention (for preventive interventions). The length of any acclimation period should be included.

2 The organism used, including the source, sequence information and passages. A statement as to

whether it is heterologous or homologous with the biological intervention.

3 The concentration of organism per unit of delivery matrix should be included, e.g. $2*10^6$ CFU per ml or per g. It is critical that the units of concentration and the delivery matrix are each specified. For organism challenges, the physiological state of the challenge organism(s) may be relevant and, as this may be influenced by the initial cultivation techniques, the details of cultivation and preparation prior to inoculation must be included.

4 Dose and route of delivery matrix administered, e.g. the challenge organisms were mixed with 100 ml whole milk administered per os.

5 The total amount of organism received, which is a function of #3 and #4. This is included as a means of checking the dose to ensure that they match.

6 The source of the isolate used in the challenge inoculum should be described, e.g. clinical isolate from a pig with diarrhoea, strain X from Y collection, N^{th} passage of virus from cell culture.

Item 5

Specific objectives and hypotheses. Clearly state primary and secondary objectives (if applicable).

Example

The primary objective of this study was to investigate if eprinomectin treatment of adult dairy cows around calving had any beneficial effects on the calving to first insemination interval, calving to conception interval, and number of inseminations per conception in herds with no or limited pasture exposure. The secondary objective was to investigate whether bulk milk ODR [optical density ratio] could be used to identify herds whose calving to conception interval could benefit from eprinomectin treatment (Sithole et al., 2006).

The objective of this study was to compare calf morbidity, mortality, and weight gain in preweaned calves reared with and without antibiotics for therapy and prophylaxis. The study hypothesis was that calf weight gain, morbidity, and mortality are not affected by antibiotics in the milk replacer or given as individual therapy (Berge et al., 2005).

Explanation

The authors should state the objectives introduced in Item 2 and the corresponding null hypothesis to be tested. Objectives (or aims) are the concepts that studies are designed to investigate. An objective usually states a broad goal to help direct the study. Hypotheses, although similar in concept, specifically state what the study is setting out to support, and allow the researcher to test a proposed hypothesis statistically. Authors should state the null hypothesis to be tested. This documents how the authors intend to achieve the objective and removes an uncertainty about the purpose of the research. Some studies are conducted to show superiority of an intervention, in which case the null hypothesis should be that the treatments are the same with respect to the primary outcome. Other studies are designed with the purpose of showing equivalence or non-inferiority of an intervention, in which case the null hypothesis is usually that the treatments are different with respect to the primary outcome (Jones et al., 1996). There is indication that although some studies are conducted with the objective of assessing equivalence, the null hypothesis is framed as for superiority studies (O'Connor et al., 2010f). By stating the null hypothesis clearly, the author will clarify the purpose of the research. This will enable the reader to interpret the meaning of non-significance correctly. This information will also allow the reader to ascertain if the sample size is correctly determined and whether the statistical methods are appropriate. If a one-tailed hypothesis test is used, then published studies justifying a unidirectional treatment effect should be referenced. If there are multiple objectives, authors should characterize them as primary versus secondary and consider ranking within the categories relative to their importance to the study's focus.

Item 6a

Clearly defined primary and secondary outcome measures.

Example

The primary outcome was IBK [infectious bovine keratoconjunctivitis] cumulative incidence over the study period. The secondary outcome was weaning weight (Funk et al., 2009).

Explanation

All trials measure at least one outcome and compare this between intervention groups. The outcomes selected for a trial need to be linked to the objectives and hypotheses. All outcomes should be identified and defined, and the methods used to measure each outcome should be described. If disease status is used as an outcome, a case definition should be provided, and person(s) responsible for assigning that diagnosis should be identified (e.g. owner/manager versus veterinarian). If specific diagnostic tests contribute to the assessment of the outcome, sensi-

tivity and specificity estimates should be included, as well as a justification of why these values are applicable to the study population. Sufficient information should be provided so that the study could be duplicated, e.g. details such as whether blood samples were collected from a coccygeal vein versus a jugular vein. If a standard approach is modified, describe the modification, rather than using phrases such as '… with slight modification'.

The primary outcome refers to the measure used to determine the study sample size (Item 7). Other outcome measures, which may be potentially equally important, but were not used to determine the sample size, should be referred to as secondary outcomes. The rationale for differentiating the outcomes as primary and secondary is to allow the reader to understand for which outcomes the study had sufficient power to detect meaningful differences in effect. In livestock trials, it is common to have one outcome related to the disease of interest (e.g. mortality or morbidity) and one related to performance (e.g. average daily gain), as these indices are often of primary concern to livestock owners. In situations where two outcomes are truly of interest and the study is designed to have sufficient power for both outcomes, the authors should provide sample size information for both outcomes (Item 7) and describe the outcome that needs the highest sample size as the primary outcome. When an outcome is measured at multiple times/points, the authors should specify which time point is the primary outcome. Secondary outcomes may also be unanticipated or unintended outcomes that become apparent as the study progressed, and it should be stated that these were unplanned outcomes.

The use of multiple outcomes is common in trials conducted in livestock populations. In a review of trials of antibiotic therapy for bovine respiratory disease, 25 of 35 studies reported multiple outcomes, and none indicated the primary outcome (O'Connor et al., 2010f). In a study evaluating reporting in food–animal trials with health or production outcomes, 91 of 100 trials reported multiple outcomes, with only four trials identifying the primary outcome (Sargeant et al., 2009a). Of 100 pre-harvest food-safety trials evaluated in a similar study, 91 reported the use of multiple outcomes, with none of the trials identifying the primary outcome (Sargeant et al., 2009b).

Item 6b

Where applicable, any methods used to enhance the quality of measurements (e.g. multiple observations and training of assessors).

Example

Corneal ulcers in the digital photographs were traced on a computer tablet (Wacom Cintiq 15X LCD tablet, Wacom Technology Corporation, Vancouver, WA, USA) using public domain image analysis software (ImageJ program; available at http://rsbweb.nih.gov/ij). Differences in magnification were accounted for by standardizing the scale of each tracing using the ruler in each photograph. The mean of three tracings of each ulcer was used to calculate the corneal ulcer surface area measurement; for data analysis, the square root of the corneal ulcer surface area was used to represent the ulcer surface area measurement (SAM). The limit of detection was 0.008 cm^2, an area corresponding to a 1-mm diameter circle. Ulcers that appeared linear or stellate were considered to be the result of mechanical trauma and were not counted as ulcers associated with IBK unless the ulcer was still present at the next weekly observation (Angelos et al., 2007).

Means of bacterial populations ($\log_{10}$ CFU/g) from each treatment were calculated from three replications for each experiment (Fabrizio and Cutter, 2005).

Explanation

Authors should provide details of any steps used to increase the precision or validity of an outcome measure. For instance, use of repeated measurements of an outcome or multiple samples may be used to define the outcome status of a study unit. The description should include the number of observations and the means of summarizing the outcome.

Standard guidelines used regarding quality of measurements should be specifically cited where relevant. Limits of detection, precision of measurements and cut-off points should always be described. When applicable, referencing validated scales and consensus guidelines is recommended to ensure transparency and reproducibility. For determination of bacterial or viral outcomes, standard procedures should be used, if available, or deviations from standard procedures should be justified. Resources for such standards are available for many areas. For example, standards for culture of mastitis pathogens in bovine milk are provided by the Clinical and Laboratory Standards Institute and National Mastitis Council (National Mastitis Council, 1987; Thompson Reuters, 2009).

Authors should provide details on any formal study-specific training of the outcome assessors, including details of inter-rater agreement during training or pre-testing. This is especially important for subjective outcomes, e.g. lameness, pain, body-condition scores and physical appearance. Many livestock studies use producer-based diagnoses of diseases, and if no additional training was provided, this should be stated.

Item 7

How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules. Sample size considerations should include sample size determinations at each level of the organizational structure and the assumptions used to account for any non-independence among groups or individuals within a group.

Examples

A sample size of 699 animals in each group was calculated to have an 80% power to detect a difference in means of 1.5 kg, assuming that the common standard deviation was 10 kg using an ANOVA with a consecutive two group *t*-test and a 5% two-sided significance level. For compensation of possible drop outs a total of 1542 healthy piglets from three consecutive farrowing batches, each comprising approximately 500 animals were included into this study (Fachinger et al., 2008).

Sample sizes were calculated by a multi-level approach with design-effects and intra-class correlations deducted from variance between [faecal] pats, groups, and farms observed in a previous field study on a similar population. The required samples sizes were 48 control farms and 48 farms in each intervention group to detect a risk ratio of 5 at 80% power with 95% confidence, when using a design effect of 13.22 to adjust for a group cluster size of 20 pat samples per group per visit. The design effect was estimated from data originating from a longitudinal study using the same sampling approach along with individual animal sampling (Ellis-Iversen et al., 2008).

Explanation

Use of an adequate sample size to detect treatment differences that are economically and biologically important is fundamental to sound trial design. The main statistical considerations in sample size calculation are the magnitude of the effect size (e.g. difference in proportions, means, survival times, etc.), standard deviation of the outcome, power ($1-\beta$ [type II error] = probability of accepting the null hypothesis when it is not true) and the significance level (α = type I error = the probability of rejecting the null hypothesis when it is true). Typically, power and significance values of 80% and 5% respectively, are used in calculations. The effect size that can be detected is inversely related to sample size – the smaller the difference, the larger the group sizes. The most common problem is lack of adequate sample size, although use of more animals than is necessary is also an important ethical concern.

For the null hypothesis and primary outcome identified in Items 5 and 6, authors should describe how the sample size was determined for each level of the organizational structure of the study setting. The description should include how non-independence of the outcome measurements and exposure were accounted for in the calculations, if relevant. If the study has multiple outcomes, and the study size chosen was considered adequate to detect clinically important differences for several outcomes, this should be reported, and the assumptions used to reach this conclusion for each outcome should be described.

Authors should state the basis for assumed values of the outcomes in the treatment groups, citing published studies whenever possible. For example, a 10% absolute difference in cumulative incidence could occur if the treated and untreated groups had incidences of 10% and 0% or 50% and 40% respectively, but the sample size required to detect the latter scenario would be greater.

In trials with long-term follow-up in production animal systems, there can be substantial loss to follow-up. For example, in a 3-year follow-up study of 100 cows in a dairy herd with 30% annual culling, only 33 of the originally enrolled cows would be expected to remain. How the anticipated loss to follow-up was accommodated should be described in later items (Items 13, 16, 20 and 21), as this may have a large effect on internal validity.

Sample size should not be confused with the specimen size. Sample size (the number of study units) and specimen size (e.g. use of 10 g of faeces versus 25 g of faeces for laboratory culture of enteric pathogens) have distinct meanings. Specimen size should be included in Item 6 (description of the outcome measures).

Example of stopping rules (from the human-health literature)

Primary end points were progression free survival, response rate, and toxicity. Overall survival was a secondary end point. Two analyses were initially planned. The first analysis was to assess and compare response rates after 21 patients were recruited to each group. If one of the groups had had a response rate less than 10% and if a difference greater than 15% in response rate was observed between the two groups, the study would have been stopped. If not, the trial could continue as a phase III study. The final planned sample size was then 91 patients in each group, on the basis of detection of a 15% difference in progression free survival between the two

arms (15% v 30% at 1 year) with a two sided test, an alpha risk of 5%, and a power of 80% (Negrier et al., 2000)

Explanation

The consensus meeting members were unaware of any livestock studies with production, health or food-safety outcomes that reported trials using stopping rules. Therefore, no examples from this literature could be provided, and the explanation for this item is quoted from the example used in the CONSORT statement elaboration document.

There are many situations where stopping rules may be applicable or useful in livestock production, and further, there are probably many published situations where authors take 'looks' at the data before the end of the study. It is not uncommon for clinical trials to recruit study units sequentially on the basis of the availability of specific inclusion criteria and in some instances, recruitment may occur over a long period of time. If an intervention is particularly efficacious, or if it causes harm, it may be ethically appropriate to end the trial early. Trials stopped early for harm should result in discontinuation or decreased use of potentially harmful interventions, and trials stopped early for benefit should contribute to earlier market availability of efficacious treatments. In the human healthcare literature, RCTs stopped early for benefit are becoming increasingly common (Montori et al., 2005). However, this decision requires that the data be examined at one or more time points during the course of the trial. This raises statistical concerns, because the multiplicity of testing increases the probability of a type I error and the identification, as significant, of random fluctuations towards greater treatment effects (Schulz and Grimes, 2005). In an example provided in the original CONSORT elaboration document, if accumulating data from a trial were examined at five interim analyses, the overall false-positive rate would be nearer to 19% than to a nominal 5% (Altman et al., 2001). Statistical methods are available for stopping procedures (Schulz and Grimes, 2005), and their use should be pre-specified in the trial protocol if interim analyses are planned. These methods generally make use of a small P-value to aid in decision making or for use as a formal stopping rule (Altman et al., 2001). The decision to stop trials early is controversial; a systematic review of trials stopped early for benefit reported implausibly large treatment effects, particularly when the number of events was small (Montori et al., 2005). An extension of this review is ongoing to further understand the extent to which trials stopped early may exaggerate treatment effects (Briel et al., 2009).

Item 8
Randomization (sequence generation). Method used to generate the random allocation sequence at the relevant level of the organizational structure, including details of any restrictions (e.g. blocking and stratification).

Example

Each heifer was randomly assigned to a treatment using a random number generator function (R Development Core Team, 2006).... (Parker et al., 2008).

Explanation

Randomization is essential to internal validity, as it is designed to minimize differences between the treatment groups and can be implemented in most RCTs, regardless of level of intervention allocation. Study units should be assigned to groups on the basis of chance (i.e. a random process), to limit the potential for confounding to influence the study result or for selection bias in the assignment of study units to treatment groups. The term 'random' has a precise meaning, wherein each study unit has a known probability of receiving a given treatment prior to assignment of the treatments. The actual treatment that a specific study unit is allocated is determined by a chance process and cannot be predicted. The methods used to generate the random allocation sequence should be reported in sufficient detail to allow the reader to assess the likelihood of bias in group assignment. Many methods of sequence generation are adequate. However, readers cannot judge the adequacy from such terms as 'random allocation', 'randomization' or 'random' without further elaboration. Therefore, authors should specify the method of sequence generation, such as a random-number table or a computerized random-number generator.

Deterministic allocation methods, such as alternate animal identification numbers, days of the week, date of birth, birth order and gate cutting, are not random (Schulz and Grimes, 2002). When these methods are used, they should not be described using the term 'random' or any variation of it. There is evidence in trials involving livestock that the term 'random' is misused to describe non-random processes, e.g. process such as gate cutting (O'Connor et al., 2010f).

When authors do not use random methods to allocate study units to treatment groups, the method of allocation should be described in a manner that would allow the reader to determine if bias was likely to be introduced because of the lack of randomization. The use of terms

such as 'systematic randomization' and 'quasi randomization' to describe these methods of allocation is not appropriate without further elaboration.

Often there are valid reasons to avoid simple randomization and to employ instead a restricted randomization method. The description of restrictors on randomization are provided in the 'CONSORT statement: Explanation and Elaboration' document and other references (Chow and Liu, 1998; Altman et al., 2001). Block randomization, also called permuted block randomization (Chow and Liu, 1998), is a method of allocation that ensures an equal distribution of study units to intervention groups and is often employed when the study size is small. The approach is to divide the whole series of study units into several blocks with equal or unequal size and randomly allocate animals to treatment within blocks, e.g. in a study of 32 animals, there may be eight blocks of four animals each. In challenge studies, which often have small study sizes, consideration should be given to employ block randomization. One disadvantage of block randomization is the potential for someone to deduce the intervention if they are aware of the block size. This risk can be mitigated by varying the block size randomly, i.e. blocks of two, four and six, within a study. An excellent description of how to implement block randomization is available (Altman and Bland, 1999). Block randomization may also be useful for field studies with group-level units of allocation, such as pen-level studies. In a pen-level study comparing two treatments with 20 pens per treatment, it may be sensible to use 10 blocks of four pens each to ensure that every group of four pens enrolled has two treated and two untreated pens (Chow and Liu, 1998).

Stratified randomization includes a covariate (thought to be a confounder) in the allocation sequence determination. For example, a feedlot study may stratify by heifers, bulls and steers, and use block randomization within each stratum to allocate to treatment group, or a swine study may control for the effect of sow and weight using stratified randomization, e.g. piglets ordered by weight from the heaviest to the lightest within a litter (a sow) and allocated to treatment in blocks of two piglets. Stratified randomization requires that block randomization be used within the strata to ensure balance of treatments within strata (Altman and Bland, 1999).

Minimization may also be used with small sample size trials to minimize differences between groups with respect to important prognostic or confounding variables. In this approach, the first study unit is assigned to treatment group using a random method; thereafter, allocation to treatment group is based on minimizing the differences among groups based on the pre-selected factor(s) (Treasure and MacRae, 1998).

Item 9
Randomization (allocation concealment). Method used to implement the random allocation sequence at the relevant level of the organizational structure (e.g. numbered containers), clarifying whether the sequence was concealed until interventions were assigned.

Example

> Sealed envelopes numbered 1 through 120 were prepared that assigned each cow to the laparoscopy-assisted abomasopexy or control group. These envelopes were opened only after confirmation of eligibility and immediately before surgery (Seeger et al., 2006).
> The remaining 57 farms were randomly allocated into three intervention groups and one control group…. The allocation was done blindly by a clerk, who assigned each participating farm a random letter drawn from an envelope, which contained one letter for each intervention group and … for the control group (Ellis-Iversen et al., 2008).

Explanation

Authors should describe whether or not any steps were taken to conceal allocation sequence until after the study unit was enrolled. The aim of allocation concealment is to prevent bias at the recruitment/enrolment phase of the trial. In a trial with adequate allocation concealment, informed consent should be obtained from the owner/manager, and the decision to include or exclude a specific study unit in the trial should be made with no knowledge of the next intervention group assignment in the allocation sequence. There is empirical evidence in the human healthcare literature that failure to report allocation concealment is associated with exaggerated treatment effects (Schulz et al., 1995; Kunz and Oxman, 1998; Moher et al., 1998; Juni et al., 2001; Kjaergard et al., 2001). Allocation concealment differs from blinding, which aims to prevent misinformation bias in the measurement of the outcome and differential management of treatment groups, and is implemented after allocation to the intervention.

An example of bias introduced as a result of failure to conceal the allocation sequence may occur in a feedlot. For example, the processing crew at a feedlot may decide not to enrol a truckload of high-risk cattle into the study if they have an unfavourable view of the treatment that truckload will be allocated. Concealment of the treatment to be received until after the cattle have been enrolled would prevent the introduction of such a bias. Similarly,

in a dairy cattle study, the owner/manager may wish certain cattle to be assigned to the treatment group because of their genetic value or severity of disease. If the person implementing the allocation sequence is unaware of the next assignment, the person is not able to be consciously or unconsciously influenced by the owner's preference. Currently, it is not common for livestock studies to use formal allocation concealment. However, inadequate allocation concealment can subvert the random allocation process (Schulz and Grimes, 2002).

Item 10
Randomization (*implementation*)**.** Who generated the allocation sequence, who enrolled study units and who assigned study units to their groups at the relevant level of the organizational structure.

Example

> Prior to the ISU [Iowa State University] farm visits, containers holding the autogenous vaccine and placebo were re-labeled injection A or B by staff who would not enroll animals at the farm. A chute processing order sheet was created, and a corresponding random allocation number between 0 and 1 was generated by an investigator not involved with enrollment (Excel, Microsoft, Redmond, WA)... At the ISU farm, three students, including the 1st author allocated the animals to treatment cohorts (Funk et al., 2009).

Explanation

For the reader to evaluate allocation concealment, it is necessary to know who generated the allocation sequence, who enrolled study units into the trial and how study units were allocated to the treatment groups. Ideally, the person(s) who generated the random allocation sequence should not be involved in the enrolment and assignment of study units to the treatment groups, as this could result in bias. In the human healthcare literature, the concern of not separating allocation generation from implementation is that if the person who generated the allocation sequence is the same person who enrols participants or assigns treatment, knowledge of the allocation sequence could influence them when interviewing potential trial participants (Schulz and Grimes, 2002). In trials in livestock populations, this bias could occur when selecting study units for participation, or could be inadvertently introduced when communicating with owners/managers on potential study units for inclusion. In some instances, owners/managers may be the person(s) enrolling study units, in which

case they should not be aware of the allocation sequence, for the same reasons.

Item 11
Whether or not those administering the interventions, caregivers and those assessing the outcomes were blinded to group assignment. If performed, how the success of blinding was evaluated. Provide justification for not using blinding if it was not used.

Example

> Two bottles, labeled 'A' and 'B', were provided to each feedlot, so that the feedlot personnel were blind to the status of the vaccine. One bottle held the vaccine... The other bottle held the placebo, which was the same as the vaccine but without the antigen (VanDonkersgoed et al., 2005).

Explanation

In controlled trials, blinding (synonym: masking) refers to the process of keeping different individuals involved in the trial unaware of the group allocation. Blinding is associated with internal validity and can be implemented in most RCTs, regardless of the level of intervention allocation. Often, the use of blinding is reported poorly in livestock trials; only four of 100 randomly selected livestock trials with health or production outcomes, and zero of 100 randomly selected pre-harvest food-safety trials reported blinding of the person administering the treatment and blinding of the outcome assessor (Sargeant et al., 2009a,b).

Trials which failed to report blinding and randomization in a systematic review of vaccines to prevent pink-eye in cattle were more likely to report favourable outcomes compared with trials that did report randomization and blinding (47% versus 20%) (Burns and O'Connor, 2008). This is consistent with studies in the human health literature that have observed larger treatment effects in trials not reporting the use of blinding (Schulz et al., 1995; Juni et al., 2001; Kjaergard et al., 2001).

It is insufficient to state that 'staff were blinded to intervention groups'; the process of achieving blinding should be reported. As with allocation, the method of blinding should be described to allow the reader to assess the validity of the blinding. The terms 'single-, double- and triple-blinded' may be used to describe the blinding, but such terms are ambiguous; a study in the human healthcare literature illustrated that individuals may have different interpretations of who is blinded when these terms are used (Devereaux et al., 2001). In addition, study subjects in animal studies cannot be blinded, unlike

human study subjects. Therefore, it is preferable to state which individuals were blinded. In livestock studies involving production, health and food-safety outcomes, we propose that authors address three potential levels of blinding: individuals associated with assessment of the outcome, individuals caring for the animals and data analysts. Individuals associated with assessment of the outcome may include owners/managers, animal caregivers, data collectors and assessors of outcomes (Devereaux et al., 2005). The personnel who are blinded should be explicitly described in this item and their role in the study should be defined (e.g. veterinarians, data analysts and personnel in laboratories performing tests).

The rationale for blinding individuals responsible for assessing the outcome is to prevent introduction of information bias. If the assessor is aware of the groups, they may over- or underestimate the outcome. Even objective outcomes such as weight gain may be biased by the lack of blinding. For example, in a study evaluating the impact of an intervention on the presence of *Salmonella* spp. on poultry carcasses at an abattoir, laboratory staff may re-examine plates more frequently, looking for *Salmonella* spp., if they are aware that a set of plates is associated with a particular intervention group expected to have higher *Salmonella* recovery rates. The CONSORT explanation and elaboration refers to this as ascertainment bias.

Further, it is also critical that anyone responsible for animal-care decisions is unaware of the group allocation. Knowledge of the intervention by caregivers may lead to differential care of the groups, which may introduce performance bias. For example, a challenge trial may be designed to assess the impact of a vaccine on the presence of clinical signs of respiratory disease after challenge. The study protocol may include a blinded person responsible for allocation of the intervention (described in Item 9), an unblinded caregiver and a blinded outcome assessor. The primary and secondary outcomes of interest may be the presence of sneezing and coughing at a certain time of day, and 21-day weight gain respectively. This study protocol may not prevent the introduction of bias if the unblinded caregiver increases observations of a particular intervention group and administers antibiotics to animals in the group at an earlier stage of disease. Increased administration of antibiotics may affect the prevalence of clinical signs and weight gain in that group, thus introducing a bias in both outcomes, although the outcome assessor is blind to the group allocation.

It is not always possible to use blinding, for example, if the intervention is a comparison between a surgical treatment and a medical treatment. In challenge studies, it may be difficult to maintain blinding if challenged animals become morbid and there is a pronounced treatment effect.

If the study cannot be blinded, authors should describe why it was not and how the study was adapted to eliminate selection and/or information bias. This should include the use of at least one objectively measured outcome.

Item 12a
Statistical methods used to compare groups for all outcome(s). Clearly state the level of statistical analysis and methods used to account for the organizational structure, where applicable.

Example

The experimental unit used for statistical analyses was individual mammary quarter. Generalized linear mixed models were used to examine risk factors for development of a new IMI [intramammary infection]. Specialized statistical techniques were used to account for clustering of quarters within cows and for clustering of cows within herds [referenced in original article]. It was assumed that the degree of similarity between observations within a cluster was the same for all clusters. The main predictor of interest was treatment, and models with the following outcomes were analyzed: new IMI caused by any pathogen, new major IMI, new environmental IMI, new gram-negative IMI, and new streptococcal IMI. For each outcome, a single model that incorporated terms for group and treatment within group was constructed (Sanford et al., 2006).

Explanation

A complete and accurate description of statistical analyses allows the reader to assess the validity of the statistical methods and the likelihood that analytical bias affected the internal validity of the study. The statistical analysis of RCT data should follow logically from the design of the study. Particular care is needed for analysing data from a trial where the units of allocation and outcome measurement are not the same. Ignoring differences between the unit of allocation and the unit of outcome measurement may lead to spurious results (Donner and Klar, 2004; Campbell et al., 2007; St-Pierre, 2007). It is critical that authors clearly describe the statistical approach to analysis employed to account for such a design. Several statistical methods of data analysis may be suitable, depending on whether the outcome measurement is continuous, ordinal or binary. There are many useful publications that appropriately describe the statistical methods to use, and consultation with a statistician in the design and analysis stage of a clinical trial is strongly recommended (St-Pierre, 2007). Further, authors are

encouraged to consult texts that describe how to write about statistical methods, as the following notes do not cover all possible contingencies (Miller, 2005).

The statistical procedure to analyse each outcome should be explicitly described. Authors should report underlying assumptions associated with each analysis (e.g. normally distributed data) and, when conducted, data transformations should be stated and justified.

The assumption of independence and identical distribution is commonly violated in livestock studies when there are multiple repeated observations per study unit over time and/or when study units are aggregated in groups and the outcomes of multiple groups are considered in the analysis. Therefore, independence and identical distribution should be considered and, where necessary, clearly described and justified. Treating each observation as an independent event when the organizational structure of the population implies non-independence is a serious violation of inherent assumptions of many statistical tests and usually leads to an overly optimistic P-value (the probability of observing the data or a more extreme result when there is no treatment effect). The statistical approach used to account for non-independence should be clearly described. An extension of the CONSORT statement for clustered trials has been developed and provides recommendations for reporting this type of trial (Campbell et al., 2004b).

Authors should provide details of all descriptive and hypothesis testing analyses that were conducted, including the name of the test used, such as t-test, chi-square test for proportions, Fisher's exact test, Mann–Whitney test or others. If the method is novel, a reference for the approach should be provided. If logistic regression modelling is used, the level of the outcome being modelled should be described, for example, 'we modelled the probability of being disease positive'. For all models, authors should indicate the data form (e.g. continuous or categorical) for all variables in the model. For categorical intervention variables, the referent must be clearly stated, for example, 'the referent level of the intervention was Treatment A'. Guidelines for reporting regression models are available (Ottenbacher et al., 2004; Tetrault et al., 2008).

Item 12b
Methods for additional analyses, such as subgroup analyses and adjusted analyses.

Example

Bacterial and clinical cures among groups were compared by chi-square tests. A stratified analysis of treatment effects was then performed to compare these effects with those after stratification on farm

(three levels) and pretreatment bacterial isolates (four levels). These analyses determined whether the treatment effects were independent of farm and primary bacterial isolate (Guterbock et al., 1993).

Explanation

In RCTs, randomization should limit the impact of confounding on the study outcome. Therefore, there is generally no need to adjust for confounding. Further, adjustment for statistically significant baseline differences is not recommended (Oxman and Guyatt, 1992; Brookes et al., 2001, 2004; Hernandez et al., 2006; Wang et al., 2007). Therefore, if authors wish to explore confounding using multivariate analysis, the rationale for assessment of confounding should be provided. Although confounding by important prognostic variables should be removed through randomization of treatments, it may still be of interest to a researcher to investigate interactions between the treatment groups and important covariates. If the interactions are significant, it may be necessary to conduct subgroup (or strata-specific) analyses. If subgroup analyses are used, the method should be clearly described. However, *post hoc* subgroup analysis is discouraged, as these comparisons may result in spurious results by increasing the number of comparisons evaluated, and the sample size is generally calculated on the basis of the full sample rather than the sample size provided by a subgroup. Therefore, subgroup analyses generally do not have credibility, and their findings are often not confirmed by subsequent studies.

Subgroup analysis often employs multivariate regression models with interaction (cross-product) terms to assess the presence of effect modification. If regression modelling is used, the authors should describe the test used to assess the significance of the interaction term. Further, the outcome being modelled, the variables of interest and covariates included in the model should be clearly stated. For all models, authors should indicate the data form, continuous or categorical, for all variables in the model. For categorical variables, the referent should be identified. Authors are encouraged to refer to guidelines for reporting regression models (Ottenbacher et al., 2004; Tetrault et al., 2008). This information is necessary to allow the reader to assess the validity of the adjusted or subgroup analyses and the likelihood of analytical bias.

Results

Item 13a
Flow of study units through each stage for each level of the organizational structure of the study (a diagram is

strongly recommended). Specifically, for each group, report the number of study units randomly assigned, receiving intended treatment, completing the study protocol and analysed for the primary outcome.

Example

Of the 939 cows (3,731 mammary quarters) enrolled in the study, 519 were assigned to group 1 (results of bacteriologic culture of all 4 quarter milk samples collected 14 days prior to the end of lactation were negative) and 420 were assigned to group 2 (results of bacteriologic culture of 1 or more quarter milk samples collected 14 days prior to the end of lactation were positive). However, 111 cows in group 1 were excluded for the following reasons: abortion ($n = 10$), disease (1), death (7), removal from the herd (2), … Similarly, 93 cows in group 2 were excluded for the following reasons: abortion ($n = 3$), death (4), removal from the herd (3), 1 or more milk samples were not collected (12), 1 or more milk samples were lost (1), the incorrect treatment was given (1), the nonlactating period lasted <30 days (6), …. Thus,

data from 734 cows (408 assigned to group 1 and 326 assigned to group 2) and 2,771 quarters were included in analyses (Sanford et al., 2006).

Twenty-eight of the 30 cows with LDA were successfully surgically treated with omentopexy via right flank laparotomy or 2-step laparoscopy-guided abomasopexy and discharged from the hospital. One cow in each surgery group died or was euthanatized (both at day 7 after surgery) because of failure to respond to treatment and subsequent multiorgan failure. Necropsy revealed extensive hepatic lipidosis in both cows, and data from both were included in the statistical comparison (Wittek et al., 2009).

Explanation

Authors should include the organizational levels applicable to their trial. Table 2 contains a list modified from the 'CONSORT: Explanation and Elaboration' document with the details required to chart the progress of owners/managers and study units through an RCT. For example, if the study solicited participation from randomly selected

Table 2. Information required to document the flow of participants through each stage of a randomized controlled trial

Stage	No. included	No. not included/excluded	Rationale
Enrolment	Owners/managers evaluated for potential enrolment	Owners/managers who did not meet the inclusion criteria Owners/managers who met the inclusion criteria, but declined to be enrolled	This information aids in determining whether animal owner/managers were likely to be representative of all owners/managers with similar livestock operations; it is relevant to assessment of external validity
	Herds/sites/pens/animals evaluated for potential enrolment	Proportion of herds/sites/pens/animals meeting inclusion criteria but not enrolled (at each level of organization)	This information aids in determining whether the enrolled number (the sample population) represents a large component of the potential study population within the facility; it is relevant to assessment of external validity
Randomization	Study units randomly assigned	May need to be described at more than one level of organization (e.g. animals randomly assigned to pens, pens randomly assigned to treatments)	Crucial for defining trial size and assessing whether a trial has been analysed by intention to treat
Treatment allocation	Study units that received treatment as allocated, by study group	Study units that did not receive treatment as allocated, by study group	Important for assessment of internal validity and interpretation of results
Follow-up	Study units that completed intervention protocol as allocated, by study group	Study units that did not complete intervention protocol allocated, by study group	Important for assessment of internal validity and interpretation of results. May also provide information about the feasibility of the protocol
Follow-up	Study units that received the full intervention protocol and completed follow-up as planned, by study group	Study units that received the full intervention protocol by study group, but did not complete follow-up as planned	Important for assessment of internal validity and interpretation of results. May also provide information about the feasibility of the protocol
Analysis	Study units included in main analysis, by study group	Study units excluded from main analysis, by study group	Crucial for assessing whether a trial has been analysed by intention to treat; reasons for excluding participants should be given

farms identified in a county-level database, then the number of farms that refused to participate should be reported. For a trial conducted on a single feedlot, which was selected by convenience, the narration/flow chart might begin with a discussion of the feedlot pens selected from within the feedlot to be included in the study.

Flow of study units in challenge trials of short duration with no losses, no protocol failure and no change in organizational structure to report may be reported effectively in the text. However, for more complex trials, authors should strongly consider including a flow chart of the trial. The complexity of the organizational structure is important to understand the external validity, whereas loss to follow-up and protocol failures affect internal validity. Thus, the reader needs this information to assess the validity of the study. For example, loss to follow-up of barns from one production system or site may have different implications than exclusion of barns distributed across multiple production companies or sites. Likewise, in livestock-production operations, animals might be sold before outcomes are assessed and, if the sale was associated with the outcome, this might result in biased results. The reader may find a chart depicting these changes easier to follow than reading a description of events. The description of losses to follow-up or protocol deviations should clearly identify these features at both the level of treatment allocation and the level of outcome measurement.

An example of loss to follow-up in a livestock study could be a study that assessed the impact of antibiotics on weight gain in the first 21 days post-arrival. For cattle that die prior to the end of the study period, weight gain cannot be assessed; therefore, these study units are lost to follow-up. In the same study, it is possible that one or more farms could decide to terminate their involvement prior to the end of the trial. In this example, the number of farms, pens and animals that were lost to follow-up should be described. In a food-safety trial assessing the impact of a vaccine on *E. coli* O157 levels in carcasses, carcasses eliminated during processing because of condemnation cannot be assessed for *E. coli* O157 status and are an illustration of study units lost to follow-up. Samples that are collected but subsequently lost in transit, or have illegible labels preventing accurate identification, are also examples of follow-up losses.

Item 13b
Describe protocol deviations from study as planned, together with reasons.

Examples

A random binary selection process was used to determine which pen of each pair received vaccine

product, except that in a few circumstances one pen of a pair had already received its arrival processing before enrollment in the study; therefore, the other pen received the vaccine (Smith et al., 2009).

Choice of surgical technique was assigned systematically to 1 of 2 groups in alternating sequence when the situation permitted. However, because the study was conducted on farms, choice of technique was often influenced by factors such as needs of the producer, availability of laparoscopic instruments, or constraints of the teaching environment (Roy et al., 2008).

Explanation

Any deviations from the trial protocol as defined prior to the start of a trial should be described; if no deviations occurred, this should also be clearly stated. Types of deviations that should be described include unplanned changes in the intervention(s), as well as changes to the way in which data were collected or analysed. If a flow diagram was used to describe participant numbers at each stage of the trial (Item 13a), it may be possible to detail some or all of the protocol deviations in this diagram. In particular, if the trial is not being conducted under the 'intention-to-treat' principle, the flow diagram can be used to indicate the exclusion of study units that were not found to meet eligibility criteria (Item 16) post-randomization. However, merely stating that a deviation occurred is not enough to justify post-randomization exclusion – details of the deviation and the reasons for the exclusion must both be provided. The number of study units that withdrew prior to collection of outcome data should also be described; if outcome data are collected for all enrolled study units, this should be stated.

Item 14
Dates defining the periods of recruitment and follow-up.

Example

Of 437 cows (1748 quarters) initially enrolled at dry off between March 27, 2002, and August 1, 2002, 419 cows remained in the study, calving between May 11, 2002, and October 5, 2002 (Godden et al., 2003).

Explanation

Knowledge of the time period during which a study took place and over what period study units were evaluated place the study in historical context (Moher et al., 2001d). Animal studies, especially those that are conducted outdoors under field conditions, may be influenced by seasonal and

Table 3. [Item 15 Example] Baseline characteristics of treatment cohorts in a randomized field trial comparing an autogenous vaccine to a placebo vaccine on three university-owned beef cattle farms[a]

| | Iowa State University | | University of Wisconsin | | | |
| | | | Farm no. 1 | | Farm no. 2 | |
	Vaccinated (n = 105)	Unvaccinated (n = 109)	Vaccinated (n = 38)	Unvaccinated (n = 37)	Vaccinated (n = 38)	Unvaccinated (n = 38)
Enrollment weight (kg) (mean ± SD)	77 (±19)	78 (±17)	110 (±18)	107.3 (±20)	80 (±14)	80 (±13)
Parity (%)						
1–3	55 (52.2)	65 (59.6)	22 (57.9)	26 (70.3)	12 (31.6)	12 (31.6)
>3	50 (47.8)	44 (40.4)	16 (42.1)	11 (29.7)	26 (68.4)	26 (68.4)
Sex (%)						
Heifer	49 (46.6)	53 (48.6)	19 (50)	24 (64.9)	18 (47.4)	22 (57.9)
Bull	56 (53.4)	56 (51.4)	19 (50)	13 (35.1)	20 (52.6)	16 (42.1)

[a]Reproduced with permission (Funk et al., 2009, p. 4588).

related weather effects. In addition, unusual weather conditions, such as extremes in temperature, drought or excessive rain or snow, may also influence the results. The length of the study should be included, and conditions which may be unique to one group should be noted, although a parallel design should avoid this issue. If a study is conducted where the control and intervention groups start and end on different dates, then this should be noted in the report.

Item 15

Baseline demographic and clinical characteristics of each group, explicitly providing information for each relevant level of the organizational structure. Data should be reported in such a way that secondary analysis, such as risk assessment, is possible.

Example

[Table 3.]

Explanation

The aim of reporting baseline information is to summarize the actual characteristics of the study population. It is important for those reading the trials to know the characteristics of the study units included in the trial, to evaluate the internal and external validity of the trial results. Providing information on whether the treatment groups were comparable with respect to important demographic and clinical characteristics allows the reader to assess the comparability of groups. Therefore, for each group, report important characteristics of study units at all relevant organizational levels. For example, baseline demographics may include herd-level characteristics such as farm size, stocking density and geographical location,

whereas animal-level demographic variables may include weight, or age and sex.

Randomized controlled trials aim to compare groups of 'study units' that differ only with respect to the intervention (treatment). Although formal random assignment to treatment groups should prevent selection bias, it does not guarantee that the groups are equivalent at baseline. However, any differences in baseline characteristics after randomization are the result of chance rather than bias (Altman and Dore, 1990). Conducting and reporting significance tests of baseline differences are not warranted (Altman and Dore, 1990; Schulz et al., 1994; Senn, 1995) and adjustment for variables on the basis of statistically significant differences at baseline is likely to bias the estimated treatment effect.

Baseline information is often efficiently presented in a table. For continuous variables, such as weight or blood pressure, the variability of the data should be reported, along with average values. Continuous variables can be summarized for each group by the mean and standard deviation. When continuous data have an asymmetrical distribution, a preferable approach may be to quote the median and a percentile range (e.g. the 25th and 75th percentiles) (Altman et al., 1983). Standard errors and confidence intervals are not appropriate for describing variability; they are inferential rather than descriptive statistics. Variables making up a small number of ordered categories (such as stages of disease I–IV) should not be treated as continuous variables; instead, numbers and proportions should be reported for each category (Altman et al., 1983; Lang and Secic, 1997).

Item 16

Number of study units (denominator) in each group included in each analysis and whether the analysis was by

'intention-to-treat'. State the results in absolute numbers when feasible (e.g. 10/20, not 50%).

Example

The surgical procedure was successfully completed in 59 of 60 (98.3%) cows in the laparoscopy-assisted abomasopexy group and 60 of 60 (100%) cows in the omentopexy (control) group. In the 1 cow in which we were not able to successfully complete the surgical procedure, extensive adhesion of the abomasum to the left ventral abdominal wall resulted from a perforating ulcer, and repositioning was therefore not possible. That cow was euthanatized and the diagnosis confirmed during necropsy. Thus, data for that cow were excluded from further evaluation (Seeger et al., 2006).

The analyses were conducted on 1367 pigs born alive that were nursing 126 sows. The standard care study group involved 60 litters with 647 piglets born alive, while the maximal care study group contained 66 litters with 720 piglets born alive. One maximal care sow was removed from the analysis because she was suspected of having clinical porcine reproductive and respiratory syndrome virus (PRRS) because all of her pigs were born weak and she was anorexic. Another sow in the standard care group was removed due to savaging. All pigs nursing these sows were removed from the study. In addition, 107 pigs died before reaching 16 d of age, and so could not be included in the analysis of the 16-d BW (Dewey et al., 2008).

In order to evaluate the measure of effect, univariable as treated analysis (AT) and intention to treat (IT) was used (Ellis-Iversen et al., 2008).

Explanation

The number of study units analysed in each intervention group for each outcome is critical for understanding the internal validity of the study. This information allows the reader to assess loss to follow-up and protocol deviations for all outcomes, as Item 13 addressed only the primary outcome. Presenting the number of participants for binary outcomes is important, because the event frequency should be taken into account when interpreting effect measures such as the risk ratio.

Intention-to-treat analysis relates to the treatment of study units that have completed the study; therefore, an outcome is available. Intention-to-treat analysis means that study units are maintained in their allocated group regardless of any protocol deviations, and that randomization is preserved. Protocol violations commonly occur when animal caregivers deviate from the protocol. For example, a caregiver may decide to add an additional antibiotic if they believe the animal is not responding to the randomly assigned treatment. A protocol violation may also occur if a poultry carcass is deemed eligible for inclusion in a processing-level trial, but is sent for re-processing and thereby not available for sampling as part of the regular processing system.

Intention-to-treat analysis represents the combined effect of the application of the protocol as well as the protocol itself, and may yield different results from analyses that only include 'per-protocol' study units. Inclusion of intention-to-treat and per-protocol analysis is strongly recommended when assessing protocols that involve changing management practices. Different outcomes from intention-to-treat and per-protocol analysis may suggest problems with the implementation of the management practices rather than the actual practices. For example, in a study on the impact of biosecurity practices on disease rates on swine farms, some farms may not conscientiously apply the biosecurity practices and violate the assigned protocol. In this instance, and assuming that biosecurity does reduce disease rates, the intention-to-treat analysis would have a smaller treatment effect than the 'per-protocol' analysis. The difference between the two analyses would have resulted from compliance issues rather than biological efficacy of the biosecurity practice *per se*. Understanding the nature of protocol violations may be valuable to future research and recommendations. Intention-to-treat analysis might suggest that the proposed protocol is not effective; however, subsequent per-protocol analysis may suggest that, when consistently applied, the practices are efficacious. Such information is useful in designing further producer programmes.

The difference between per-protocol and intention-to-treat analysis may not be applicable when the opportunity for protocol violation is rare, as occurs with challenge studies of short duration which involve a onetime application of the intervention.

Item 17
For each primary and secondary outcome, a summary of results for each group, accounting for each relevant level of the organizational structure and the estimated effect size and its precision (e.g. 95% confidence interval).

Example
[Table 4.]

Explanation

For each primary and secondary outcome reported in Item 6, a summary outcome should be reported for each

intervention group. The rationale for providing this information is to allow the reader to assess the clinical relevance as well as the statistical significance of the differences between the interventions groups, information that is better conveyed by summary effect measures rather than by the sole use of *P*-values. Further, as meta-analyses and stochastic modelling are sometimes conducted several years after primary studies are reported, it is also advisable to provide raw summary data for all relevant subpopulations.

For continuous outcomes, the mean and standard deviation should be reported with the number in each group, rather than reporting the mean difference. When reporting proportions from binary data, include the absolute numbers as well as the percentage or proportion (10/20 combined with 50% or 0.5%).

A contrast measure ('effect measure') between the groups should also be included. For binary outcomes, this may be the rate ratio, risk ratio, odds ratio, rate difference or risk difference. For survival data, the most commonly used effect measure is the hazard ratio. For continuous data, the effect measure generally is the differences in mean values among intervention groups. For each effect measure, the 95% confidence interval should be reported. If authors wish to include the *P*-value, it should be in addition to, not a substitute for, the 95% confidence interval. Confidence intervals convey considerably more information than *P*-values, and are preferred (Gardner and Altman, 1986).

It should be clear whether the effect measure was unadjusted (i.e. a bivariable comparison between the intervention groups) or whether it was adjusted for confounding variables (not encouraged), non-independence or both. Given the impact of the extent of the intra-cluster correlation on the power of the study, the intra-cluster correlation coefficient or k statistic for each outcome should also be provided (Donner, 2000). When interaction is present, effect measures for each level of the interacting variable should be reported.

Results should be reported for all planned analyses, including those that did not find a statistically significant association between the intervention and the outcome. If the study was conducted at multiple sites, site-specific summary information should be provided as well as overall summary information. This will allow readers to assess variation in the effect measure across sites.

It is not recommended to report the parameter estimates for logistic or Poisson models, as it is unnecessary work for the reader to convert the parameter estimate to an effect measure. In addition, it may not be possible to calculate the effect measure if the authors failed to specify whether a deviation from the mean versus reference coding was used in the modelling described under Item 12.

Item 18

Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those pre-specified and those exploratory.

Example

The rate of bacterial cures did not differ ($P = 0.61$) between oxytocin-treated and antibiotic-treated cows (Table 3 [in original citation]). Clinical cure rates were nearly identical ($P = 0.99$) for the three treatment groups. Treatment did not significantly influence clinical or bacterial cure rate when the data were stratified by herd ($P = 0.27$) (Table 3). When the data were stratified by organism isolated, bacterial cure rate did not differ by treatment (Table 4 [in original citation]). Clinical cure rate did not differ by treatment, except that treatment with either antibiotic improved clinical cure rate ($P = 0.02$) for the category of other bacteria (Table 4) (Guterbock et al., 1993).

Explanation

As the number of analyses using the same data increases, so does the risk of false-positive findings

Table 4. [Item 17 example] Effect of tilmicosin (MIC) and tulathromycin (DRAX) on feedlot performance of feedlot heifer calves at moderate risk for bovine respiratory disease[a]

Performance variable	Experimental group			
	MIC	DRAX	SEM	*P*-value
No. of pens	10	10		
No. of heifers	2250	2244		
Processing weight (lb)	604	603	0.90	0.70
DOF at terminal implant	137	137	0.14	0.34
Terminal implant weight (lb)	1015	1024	2.48	0.03*
DDMI at implant (lb)	18.7	19.2	0.14	0.03*
ADG at implant (lb/day)	3.02	3.09	0.02	0.03*
DMC at implant (lb/lb)	6.57	6.50	0.06	0.38
DOF at harvest	218	218	–	1.0
Final weight + (lb)	1243	1244	2.35	0.86
Final weight − (lb)	1246	1246	2.38	0.99
Final DDMI (lb)	20.0	20.3	0.18	0.28
Final ADG + (lb/day)	2.92	2.9	0.008	0.28
Final ADG − (lb/day)	2.87	2.87	0.009	0.99
Final DMC + (lb/lb)	6.87	6.95	0.06	0.32
Final DMC − (lb/lb)	6.97	7.02	0.05	0.49

[a]Reproduced with permission (VanDonkersgoed et al., 2008, p. 293)
*Statistically significant differences ($P \leq 0.05$).
+, weight of dead animals added; −, weight of dead animals removed; ADG, average daily gain; DDMI, daily dry matter intake; DMC, dry matter conversion; DOF, days on feed.

(Tukey, 1977). Examples of multiple analyses include evaluating the intervention against multiple outcomes; performing multiple analyses based on control of different potential confounding variables or within levels of an interaction variable and subgroup analysis, wherein interventions are evaluated within a subgroup of study units on the basis of an important characteristic (e.g. age group). Multiple outcomes are commonly used in published trials. An evaluation of 100 livestock health-and-production trials and 100 pre-harvest food-safety trials reported a mean number of outcomes per trial of 9.5 (range, 1–41) and 8.5 (range, 1–51) respectively (Sargeant et al., 2009a,b). In trials with large numbers of outcomes, the risk of a type I error is substantial, and significant associations may be over interpreted. Studies with multiple outcomes and/or subgroup analyses also have a high risk of a type II error, as the power of the study is usually calculated for the primary comparison and not for additional analyses.

As discussed, subgroup analysis may be planned and described *a priori* (preferred) or may be included as a *post hoc* decision on the basis of preliminary analyses. If the latter is the case, the *post hoc* nature of the decision should be clearly stated, and the results of the subgroup analysis should be described as exploratory. Experience from human healthcare suggests that authors should resist the temptation to perform post hoc subgroup analyses (Yusuf et al., 1991; Oxman and Guyatt, 1992; Assmann et al., 2000; Brookes et al., 2004; Lagakos, 2006). Analyses that were pre-specified in the trial protocol are much more reliable than those suggested by the data. Authors should already have indicated which analyses were pre-specified in Items 2, 5, 6 and 12.

When subgroup or adjusted analyses are performed, information should already have been provided on the specific subgroups that were analysed and the reasons for such analyses (see Item 12). All subgroup analyses that were performed should be reported, regardless of the results; bias may result from selective reporting of subgroup analyses. Results from any formal tests of interaction (Item 12b) should be provided in terms of estimated differences in the intervention effect in each subgroup, including a confidence interval, rather than only a *P*-value. A recent study reported that 59 of 97 trials involved subgroup analyses, but only 46% reported interaction tests for some or all subgroup analyses (Wang et al., 2007). Another study, involving 63 RCTs, found that only 11 of 39 RCTs with subgroups included tests of interaction (Hernandez et al., 2006). Additionally, details on analyses and justifications for analyses should be provided whenever adjustments are made for baseline variables. If the study included such adjustments, authors should specify whether the adjustments and selection of adjusted variable(s) were planned. Both unadjusted and adjusted result should also be provided.

Item 19
All important adverse events or side effects in each intervention group.

Examples

Postsurgical complications were observed in 7 (11.6%) cows of the abomasopexy group, which did not differ significantly ($P = 0.163$; Fisher exact test [2-sided]) from the number of cows with postsurgical complications (2 [3.3%]) in the control group. Two cows in the abomasopexy group developed moderate localized peritonitis that was more severe than expected after the surgical procedure. Peritonitis was diagnosed on the basis of clinical signs (fever, tenseness of the abdominal wall, and moderate decrease in general condition) and results of transabdominal ultrasonography. Furthermore, three cows developed cellulitis at the abomasopexy site, which was recognizable as a phlegmonous swelling of the abdominal wall, and 2 cows had a relapse of the LDA after they had kicked the gauze bandage off. For both cows with relapse, a second laparoscopy-assisted abomasopexy was successfully performed. None of the cows in the control group had relapse of the LDA, but two cows developed a purulent infection at the omentopexy site. All wound infections (three cows with cellulitis in the abomasopexy group and two cows with purulent infection in the control group) resolved after parenteral administration of an antimicrobial for several days (Seeger et al., 2006).

Two hundred and sixty-six animals were allocated to the LA 30 group, 265 animals were allocated to the LA 20 group, and 266 animals were allocated to the FLOR group. There were no adverse reactions in any of the experimental groups (Schunicht et al., 2002).

Explanation

Many interventions have unintended and often undesirable effects in addition to intended effects. Readers need information about the harms as well as the benefits of interventions to make rational and balanced decisions. The existence and nature of adverse effects can have a major impact on whether a particular intervention will be deemed acceptable and useful. In livestock studies, adverse reactions would include any occurrence that may affect animal health, appearance or performance. Further, an adverse event may include reduced meat quality or

safety. For example, studies of management practices during transportation may observe negative impacts on the carcass grade or increased condemnations, and such adverse events should be reported.

Not all reported adverse events observed during a trial are necessarily a consequence of the intervention; some may be a consequence of the condition being treated. Randomized controlled trials offer the best approach for providing safety data as well as efficacy data, although they cannot be relied upon to detect rare adverse effects. At a minimum, authors should provide estimates of the frequency of the main severe adverse events and reasons for treatment discontinuation separately for each intervention group. If animals experience an adverse event more than once, the data presented should refer to number of affected animals; number of adverse events may also be of interest.

Discussion

Item 20

Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes. Where relevant, a discussion of herd immunity should be included. If applicable, a discussion of the relevance of the disease challenge should be included.

Example

> The logistics of conducting research with privately owned cattle meant that we often did not know when or where pens of cattle were marketed, or we were unable to be at the site of harvest; therefore, pens of cattle were enrolled by convenience. We do not believe that this practice introduced selection bias because pens of cattle were randomly assigned to vaccine treatment initially and because research personnel were blind to laboratory results when enrolling pens for the current study. We found no evidence of selection bias based on comparing the number of cattle per pen and the number of days elapsing between arrival processing and reprocessing in this study and the larger longitudinal study from which these pens were enrolled (Smith et al., 2009).

Explanation

To encourage consistent format with the CONSORT statement, the authors of the REFLECT statement agree with the recommendation of the CONSORT Explanation and Elaboration document, which proposes that authors follow the five recommendations presented in the *Annals*

of Internal Medicine (Moher et al., 2001d): (1) brief synopsis of the key findings; (2) consideration of possible mechanisms and explanations; (3) comparison with relevant findings from other published studies (whenever possible including a systematic review combining the results of the current study with the results of all previous relevant studies); (4) limitations of this study (and methods used to minimize and compensate for those limitations) and (5) a brief section that summarizes the clinical and research implications of the work, as appropriate.

Most, if not all, trials will have some limitations. Therefore, the discussion section should include a discussion of these limitations and the possible implications they might have on the conclusions from the trial. The discussion of study limitations should include any potential biases, including the presence of uncontrolled confounding factors or differences among intervention groups (Campbell et al., 2004b), or the potential for selection bias. If possible, the impact of these potential biases should be quantified. Sensitivity analyses that illustrate the magnitude of confounding, misclassification or selection, which would be required to change the inference of the study, are preferable to statements such as 'results should be interpreted with caution because of the potential for confounding/misinformation/selection bias'. If employed, these sensitivity analyses should be described in the methods and materials, and the results sections.

If blinding or formal randomization to treatment group was not used, a discussion of the implications and objectivity of the outcome (for non-blinded studies) should be included. Authors should also discuss the number of subjects per intervention group that did not complete the study and how this may have affected the results and conclusions.

A consideration of potential imprecision in the outcome measure also may be appropriate. Imprecision may be introduced into a study at a number of points, such as when the primary outcome is measured (Item 6) or during the determination of whether a study unit meets the eligibility criteria (Item 3a). For instance, a blood test may have been validated in adult cows, but not in calves, or a laboratory technician may not be familiar with how to interpret a blood smear from a particular species. As this kind of issue has the potential to increase imprecision, such issues should be mentioned in the discussion.

Authors should address the biological and practical importance of the work carried out, while not extrapolating the results of their studies beyond the limits of their data. If the trial included the evaluation of multiple outcomes, the potential for type I errors should be discussed. Conversely, if no significant associations with the intervention were observed, authors should not interpret this as evidence of the truth of the null hypothesis.

In particular, failure to reject the null hypothesis in a superiority study should not be interpreted as evidence of equivalence (see Item 6) (Jones et al., 1996). The statistical power of the trial should already be clear from the methods and materials.

When appropriate, authors should also discuss the potential effects of herd immunity, given the study design chosen, i.e. individual or clustered allocation. An example of the possible effects of herd immunity would be the evaluation of vaccine efficacy. If a vaccine is efficacious, then one would expect the control (non-vaccinated) group also to receive some benefit because of interruption of disease transmission, if they are in contact with the vaccinates. Therefore, when animals within groups are allocated to vaccine, or when groups within a common housing area are allocated to vaccine, vaccine efficacy theoretically measures only the direct benefits of vaccination and is probably an underestimate of true vaccine efficacy. Thus, the choice of control group and the implications of that choice in terms of possible herd immunity should be discussed when applicable.

When challenge models are used, the discussion should include a consideration of the degree, to which the pathogen represents wild-type pathogens, and the dose and route of administration used in the study should be compared with the dose and route of infection occurring in a natural disease challenge.

Item 21
Generalizability (external validity) of the trial findings.

Example

Although extrapolation of results obtained in experimentally infected pigs to the field situation should be done with caution, the infection model used allows studying the effects of infections with *Mycoplasma hyopneumoniae* of different virulence in a standardized and reproducible way (Villarreal et al., 2009).

The external validity of the study may have been compromised to some extent because of the close proximity of the experimental population to the regional agricultural college (Taveros and More, 2001).

Explanation

The external validity of a study refers to the degree to which the study results can be generalized beyond the study population (Rothwell, 2005). External validity may vary, depending on the application for which the reader of the trial is considering using the intervention. Factors involved in determining external validity include the characteristics of the study units and study population, the trial setting and the interventions and the outcomes measured (Rothwell, 2005). For instance, there are often differences in the housing and management of young stock, compared with those of mature animals or animals in the finishing production stages. In ruminant animals, trials conducted in pre-weaned animals may not be relevant to post-weaned animals because of differences in nutritional physiology. Therefore, when relevant, possible limitations with extending the results of a trial to animals in different production stages should be discussed. To allow the reader to assess external validity, trial reports should include sufficient information on (1) eligibility criteria (Item 3), (2) trial setting and location (Item 3), (3) interventions and administration methods (Item 4), (4) outcome definitions (Item 6) and (5) the recruitment and follow-up periods (Item 14). However, the authors should also provide their own interpretation of the external validity of the results.

Of particular relevance to livestock production is the applicability of challenge trials. There may be substantive differences between natural and artificial disease challenges, including potential differences in the exposure dose, the strain(s) used and the route of administration. Challenge studies may also use design features such as restriction of the population and the study setting to reduce the potential for confounding to bias the outcome. However, when challenge trials are conducted in narrowly selected populations of animals, the study population may not represent the target population (e.g. on the basis of age or weight, or whether they are free of other important pathogens that might be encountered under commercial conditions). Challenge trials are often conducted in animals housed individually or in small groups in laboratory settings, which may not be representative of the environment that the target population experiences. Thus, although challenge trials may provide strong preliminary evidence of treatment efficacy, their external validity will be not as strong as that of an RCT conducted under commercial conditions.

Trials conducted in research herds also may use different pen sizes or animal densities, compared with commercial settings, and this will impact external validity. Similarly, RCTs conducted at a single commercial site may not be representative of the variety of settings possible, and the authors should acknowledge this.

In addition, information on product safety, product quality and welfare of study subjects may be useful to readers to decide on the applicability of the results.

Item 22
General interpretation of the results in the context of current evidence.

Example

In the present study, we evaluated the possible efficacy of single cow calving pens for preventing neonatal calf diseases. Utilizing single cow calving pens that are cleaned between uses did not provide added protection to calves against calf diseases. Husbandry practices other than maternity pen management could have been relatively more important determinants of preweaning health than use of single cow calving pens. While it might be true that there really is little to no added protection provided by single cow calving pens against neonatal calf diseases, cautious interpretation of the current results is in order due to lack of corroborative data since no studies had previously attempted to address similar questions using the study design employed in the present study. These findings are inconclusive (Pithua et al., 2009).

Explanation

When discussing the results of the study, the researcher should consider and include the broader impacts of the results relative to issues including, but not limited to, policy, societal welfare and concern, and industry and stakeholder concern. At a minimum, authors should discuss results of the study in the context of all previous work, regardless of whether the results are supportive or not. If authors used a Bayesian analysis, it is recommended that the description include estimates in terms of the results from previous studies. If similar studies do not exist or are not available for review, the authors should indicate this as a limitation to their results. By placing the results in the context of prior research, authors allow readers to interpret the results of available studies relative to chronological changes in animal populations (e.g. herd sizes and management practices), disease processes and interventions. Authors should avoid including *post hoc* statements about the cost benefit or cost effectiveness of the intervention unless that was a stated purpose of the manuscript, and the methods for such analyses were described in Items 6 and 12.

Acknowledgements

Grant support

Potential conflicts of interest

None disclosed.

References

Ah-See, K. W., and N. C. Molony, 1998: A qualitative assessment of randomized controlled trials in otolaryngology. *J. Laryngol. Otol.* 112, 460–463.

Altman, D. G., and J. M. Bland, 1999: How to randomise. *BMJ* 319, 703–704.

Altman, D. G., and C. J. Dore, 1990: Randomisation and baseline comparisons in clinical trials. *Lancet* 335, 149–153.

Altman, D. G., S. M. Gore, M. J. Gardner, and S. J. Pocock, 1983: Statistical guidelines for contributors to medical journals. *Br. Med. J. (Clin Res Ed)* 286, 1489–1493.

Altman, D. G., K. F. Schulz, D. Moher, M. Egger, F. Davidoff, D. Elbourne, P. C. Gotzsche, and T. Lang, 2001: The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann. Intern. Med.* 134, 663–694.

Angelos, J. A., R. G. Bonifacio, L. M. Ball, and J. F. Hess, 2007: Prevention of naturally occurring infectious bovine keratoconjunctivitis with a recombinant *Moraxella bovis* pilin-*Moraxella bovis* cytotoxin-ISCOM matrix adjuvanted vaccine. *Vet. Microbiol.* 125, 274–283.

Assmann, S., S. Pocock, L. Enos, and L. Kasten, 2000: Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 355, 1064–1069.

Begg, C., M. Cho, S. Eastwood, R. Horton, D. Moher, I. Olkin, R. Pitkin, D. Rennie, K. F. Schulz, D. Simel, and D. F. Stroup, 1996: Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 276, 637–639.

Berge, A. C., P. Lindeque, D. A. Moore, and W. M. Sischo, 2005: A clinical trial evaluating prophylactic and therapeutic antibiotic use on health and performance of preweaned calves. *J. Dairy Sci.* 88, 2166–2177.

Boutron, I., D. Moher, D. G. Altman, K. F. Schulz, and P. Ravaud, 2008: Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann. Intern. Med.* 148, 295–309.

Briel, M., M. Lane, V. M. Montori, D. Bassler, P. Glasziou, G. Malaga, E. A. Akl, I. Ferreira-Gonzalez, P. onso-Coello, G. Urrutia, R. Kunz, C. R. Culebro, S. A. da Silva, D. N. Flynn, M. B. Elamin, B. Strahm, M. H. Murad, B. Djulbegovic, N. K. Adhikari, E. J. Mills, F. Gwadry-Sridhar, H. Kirpalani, H. P. Soares, N. O. Elnour, J. J. You, P. J. Karanicolas, H. C. Bucher, J. F. Lampropulos, A. J. Nordmann, K. E. Burns, S. M. Mulla, H. Raatz, A. Sood, J. Kaur, C. R. Bankhead, R. J. Mullan, K. A. Nerenberg, P. O. Vandvik, F. Coto-Yglesias, H. Schunemann, F. Tuche, P. P. Chrispim, D. J. Cook, K. Lutz, C. M. Ribic, N. Vale, P. J. Erwin, R. Perera, Q. Zhou, D. Heels-Ansdell, T. Ramsay, S. D. Walter, and G. H. Guyatt, 2009: Stopping randomized trials early for benefit: a protocol of the Study Of Trial Policy Of Interim Truncation-2 (STOPIT-2). *Trials* 10, 49.

Brookes, S. T., E. Whitley, T. J. Peters, P. A. Mulheran, M. Egger, and S. G. Davey, 2001: Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. *Health Technol. Assess.* 5, 1–56.

Brookes, S. T., E. Whitely, M. Egger, G. D. Smith, P. A. Mulheran, and T. J. Peters, 2004: Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J. Clin. Epidemiol.* 57, 229–236.

Burns, M. J., and A. M. O'Connor, 2008: Assessment of methodological quality and sources of variation in the magnitude of vaccine efficacy: a systematic review of studies from 1960 to 2005 reporting immunization with *Moraxella bovis* vaccines in young cattle. *Vaccine* 26, 144–152.

Campbell, M. K., D. R. Elbourne, and D. G. Altman, 2004a: CONSORT statement: extension to cluster randomised trials. *BMJ* 328, 702–708.

Campbell, M. K., D. R. Elbourne, and D. G. Altman, 2005: The CONSORT statement for cluster randomised trials [spanish]. *Med. Clin. (Barc.)* 125(Suppl 1), 28–31.

Campbell, M. K., D. Elbourne, and D. G. Altman, 2006: CONSORT statement: extension to cluster randomized trials [Chinese]. *Clin. J. Evid-Based Med.* 6, 451–458.

Campbell, M. J., A. Donner, and N. Klar, 2007: Developments in cluster randomized trials and Statistics in Medicine. *Stat. Med.* 26, 2–19.

Chow, S., and J. Liu, 1998: Design and Analysis of Clinical Trials. Wiley-Interscience, Hoboken.

Cochrane, A. L., 1972: Effectiveness and Efficiency; Random Reflections on Health Services. The Nuffield Provincial Hospitals Trust, Abingdon, Berks, UK.

Cook, S. R., S. J. Bach, S. M. L. Stevenson, R. DeVinney, A. A. Frohlich, L. Fang, and T. A. McAllister, 2005: Orally administered anti-*Escherichia coli* O157:H7 chicken egg yolk antibodies reduce fecal shedding of the pathogen by ruminants. *Can. J. Anim. Sci.* 85, 291–299.

Cooke, R. F., R. N. Silva Del, D. Z. Caraviello, S. J. Bertics, M. H. Ramos, and R. R. Grummer, 2007: Supplemental choline for prevention and alleviation of fatty liver in dairy cattle. *J. Dairy Sci.* 90, 2413–2418.

Deguchi, K., E. Yokoyama, T. Honda, and K. Mizuno, 2009: Efficacy of a novel trivalent inactivated vaccine against the shedding of *Salmonella* in a chicken challenge model. *Avian Dis.* 53, 281–286.

Denagamage, T. N., A. M. O'Connor, J. M. Sargeant, A. Rajic, and J. D. McKean, 2007: Efficacy of vaccination to reduce *Salmonella* prevalence in live and slaughtered swine: a systematic review of literature from 1979 to 2007. *Foodborne Pathog. Dis* 4, 539–549.

DerSimonian, R., L. Charlette, B. McPeak, and F. Mosteller, 1982: Reporting on methods in clinical trials. *N. Engl. J. Med.* 306, 1332–1337.

Devereaux, P. J., B. J. Manns, W. A. Ghali, H. Quan, C. Lacchetti, V. M. Montori, M. Bhandari, and G. H. Guyatt, 2001: Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 285, 2000–2003.

Devereaux, P. J., M. Bhandari, V. M. Montori, B. J. Manns, W. A. Ghali, and G. H. Guyatt, 2005: Double blind, you are the weakest link – goodbye! *Equine Vet. J.* 37, 557–558.

Dewey, C. E., T. Gomes, and K. Richardson, 2008: Field trial to determine the impact of providing additional care to litters on weaning weight of pigs. *Can. J. Vet. Res.* 72, 390–395.

Dohoo, I. R., 2004: The design of randomized controlled trials of veterinary vaccines. *Anim. Health Res. Rev.* 5, 235–238.

Donner, A., and N. Klar, 2000: Design and Analysis of Cluster Randomization Trials in Health Research (Hodder Arnold Publication). Wiley, London.

Donner, A., and N. Klar, 2004: Pitfalls of and controversies in cluster randomization trials. *Am. J. Public Health* 94, 416–422.

Dubuc, J., D. DuTremblay, M. Brodeur, T. Duffield, R. Bagg, J. Baril, and L. DesCoteaux, 2009: A randomized herd-level field study of dietary interactions with monensin on milk fat percentage in dairy cows. *J. Dairy Sci.* 92, 777–781.

Elbers, A. R., and Y. H. Schukken, 1995: Critical features of veterinary field trials. *Vet. Rec.* 136, 187–192.

Ellis-Iversen, J., R. P. Smith, W. S. Van, G. A. Paiba, E. Watson, L. C. Snow, and A. J. Cook, 2008: Farm practices to control *E. coli* O157 in young cattle – a randomised controlled trial. *Vet. Res.* 39, 3.

Fabrizio, K., and C. Cutter, 2005: Application of electrolyzed oxidizing water to reduce *Listeria monocytogenes* on ready-to-eat meats. *Meat Science* 71, 327–333.

Fachinger, V., R. Bischoff, S. B. Jedidia, A. Saalmuller, and K. Elbers, 2008: The effect of vaccination against porcine circovirus type 2 in pigs suffering from porcine respiratory disease complex. *Vaccine* 26, 1488–1499.

Funk, L., M. Maroney, T. Engelton, V. Cooper, J. Kinyon, P. Plummer, and A. O'Connor, 2009: A randomized and blinded trial to assess the efficacy of an autogenous vaccine to prevent naturally occurring infectious bovine keratoconjunctivis (IBK) in beef calves. *Vaccine* 27, 4585–4590.

Gagnier, J. J., H. Boon, P. Rochon, D. Moher, and C. Bombardier, 2005: Improving the quality of reporting for randomized clinical trials evaluating herbal interventions: an extension of the CONSORT Statement [abstract]. *Clin. Trials* 2, S43.

Gagnier, J., H. Boon, P. Rochon, J. Barnes, D. Moher, and C. Bombardier, 2006a: Improving the quality of reporting of randomized controlled trials evaluating herbal interventions: implementing the CONSORT statement. *Explore (NY)* 2, 143–149.

Gagnier, J. J., H. Boon, P. Rochon, D. Moher, J. Barnes, and C. Bombardier, 2006b: Recommendations for reporting randomized controlled trials of herbal interventions: Explanation and elaboration. *J. Clin. Epidemiol.* 59, 1134–1149.

Gagnier, J. J., H. Boon, P. Rochon, D. Moher, J. Barnes, and C. Bombardier, 2006c: Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement. *Ann. Intern. Med.* 144, 364–367.

Gardner, M. J., and D. G. Altman, 1986: Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br. Med. J. (Clin Res Ed)* 292, 746–750.

Godden, S., P. Rapnicki, S. Stewart, J. Fetrow, A. Johnson, R. Bey, and R. Farnsworth, 2003: Effectiveness of an internal teat seal in the prevention of new intramammary infections during the dry and early-lactation periods in dairy cows when used with a dry cow intramammary antibiotic. *J. Dairy Sci.* 86, 3899–3911.

Gotzsche, P. C., 1989: Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control. Clin. Trials* 10, 31–56.

Guterbock, W. M., A. L. Van Eenennaam, R. J. Anderson, I. A. Gardner, J. S. Cullor, and C. A. Holmberg, 1993: Efficacy of intramammary antibiotic therapy for treatment of clinical mastitis caused by environmental pathogens. *J. Dairy Sci.* 76, 3437–3444.

Hernandez, A. V., E. Boersma, G. D. Murray, J. D. Habbema, and E. W. Steyerberg, 2006: Subgroup analyses in therapeutic cardiovascular clinical trials: are most of them misleading? *Am. Heart J.* 151, 257–264.

Higgins, A. J., 1997: Randomized controlled trials – the problem of clinical trials in veterinary science. *Vet J* 154, 1–3.

Hopewell, S., M. Clarke, D. Moher, E. Wager, P. Middleton, D. G. Altman, and K. F. Schulz, 2008: CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med.* 5, e20.

Ioannidis, J. P., S. J. Evans, P. C. Gotzsche, R. T. O'Neill, D. G. Altman, K. Schulz, and D. Moher, 2004: Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann. Intern. Med.* 141, 781–788.

Jim, G. K., C. W. Booker, P. T. Guichon, O. C. Schunicht, B. K. Wildman, J. C. Johnson, and P. W. Lockwood, 1999: A comparison of florfenicol and tilmicosin for the treatment of undifferentiated fever in feedlot calves in western Canada. *Can. Vet. J.* 40, 179–184.

Jones, B., P. Jarvis, J. A. Lewis, and A. F. Ebbutt, 1996: Trials to assess equivalence: the importance of rigorous methods. *BMJ* 313, 36–39.

Juni, P., D. G. Altman, and M. Egger, 2001: Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 323, 42–46.

Kane, R. L., J. Wang, and J. Garrard, 2007: Reporting in randomized clinical trials improved after adoption of the CONSORT statement. *J. Clin. Epidemiol.* 60, 241–249.

Kjaergard, L. L., J. Villumsen, and C. Gluud, 2001: Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann. Intern. Med.* 135, 982–989.

Kunz, R., and A. D. Oxman, 1998: The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 317, 1185–1190.

Lagakos, S. W., 2006: The challenge of subgroup analyses – reporting without distorting. *N. Engl. J. Med.* 354, 1667–1669.

Lang, T., and M. Secic, 1997: How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers. American Coll of Physicians, Philadelphia.

Lund, E. M., K. M. James, and J. D. Neaton, 1994: Clinical trial design: veterinary perspectives. *J. Vet. Intern. Med.* 8, 317–322.

Miller, J. E., 2005: Chicago Guide to Writing about Multivariate Analysis (Chicago Guides to Writing, Editing, and Publishing). University Of Chicago Press, Chicago.

Moher, D., B. Pham, A. Jones, D. J. Cook, A. R. Jadad, M. Moher, P. Tugwell, and T. P. Klassen, 1998: Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 352, 609–613.

Moher, D., A. Jones, and L. Lepage, 2001a: Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 285, 1992–1995.

Moher, D., K. F. Schulz, and D. Altman, 2001b: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 285, 1987–1991.

Moher, D., K. F. Schulz, and D. G. Altman, 2001c: The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Med. Res. Methodol.* 1, 2.

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

Moher, D., K. F. Schulz, and D. G. Altman, 2001e: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann. Intern. Med.* 134, 657–662.

Montori, V. M., P. J. Devereaux, N. K. Adhikari, K. E. Burns, C. H. Eggert, M. Briel, C. Lacchetti, T. W. Leung, E. Darling, D. M. Bryant, H. C. Bucher, H. J. Schunemann, M. O.

Meade, D. J. Cook, P. J. Erwin, A. Sood, R. Sood, B. Lo, C. A. Thompson, Q. Zhou, E. Mills, and G. H. Guyatt, 2005: Randomized trials stopped early for benefit: a systematic review. *JAMA* 294, 2203–2209.

National Mastitis Council, 1987: Laboratory Handbook on Bovine Mastitis. National Mastitis Council, Verona.

Negrier, S., A. Caty, T. Lesimple, J. Y. Douillard, B. Escudier, J. F. Rossi, P. Viens, and F. Gomez, 2000: Treatment of patients with metastatic renal carcinoma with a combination of subcutaneous interleukin-2 and interferon alfa with or without fluorouracil. Groupe Francais d'Immunotherapie, Federation Nationale des Centres de Lutte Contre le Cancer. *J. Clin. Oncol.* 18, 4009–4015.

Northcutt, J. K., J. A. Cason, D. P. Smith, R. J. Buhr, and D. L. Fletcher, 2006: Broiler carcass bacterial counts after immersion chilling using either a low or high volume of water. *Poult. Sci.* 85, 1802–1806.

O'Connor, A. M., N. G. Wellman, R. B. Evans, and D. R. Roth, 2006: A review of randomized clinical trials reporting antibiotic treatment of infectious bovine keratoconjunctivitis in cattle. *Anim. Health Res. Rev.* 7, 119–127.

O'Connor, A. M., T. Denagamage, J. M. Sargeant, A. Rajic, and J. McKean, 2008: Feeding management practices and feed characteristics associated with *Salmonella* prevalence in live and slaughtered market-weight finisher swine: a systematic review and summation of evidence from 1950 to 2005. *Prev. Vet. Med.* 87, 213–228.

O'Connor, A. M., J. M. Sargeant, I. A. Gardner, J. Dickson, and M. Torrence, 2010a: The REFLECT statement: Methods and processes of creating reporting guidelines for randomized control trials for livestock and food safety. *J. Food Prot* 73, 132–139.

O'Connor, A. M., J. M. Sargeant, I. A. Gardner, J. Dickson, and M. Torrence, 2010b: The REFLECT statement: Methods and processes of creating reporting guidelines for randomized control trials for livestock and food safety. *J. Vet. Intern. Med* 24, 57–64.

O'Connor, A. M., J. M. Sargeant, I. A. Gardner, J. Dickson, and M. Torrence, 2010c: The REFLECT statement: Methods and processes of creating reporting guidelines for randomized control trials for livestock and food safety. *J. Swine Health Prod* 18, 18–26.

O'Connor, A. M., J. M. Sargeant, I. A. Gardner, J. Dickson, and M. Torrence, 2010d: The REFLECT statement: Methods and processes of creating reporting guidelines for randomized control trials for livestock and food safety. *Zoonoses Public Health*, 57, 95–104.

O'Connor, A. M., J. M. Sargeant, I. A. Gardner, J. Dickson, and M. Torrence, 2010e: The REFLECT statement: Methods and processes of creating reporting guidelines for randomized control trials for livestock and food safety. *Prev. Vet. Med*, 93, 11–18.

O'Connor, A. M., N. G. Wellman, M. Rice, and L. Funk, 2010f: Evaluating the Quality of Reporting Randomized Clinical Trials for Antibiotic Treatment of Bovine Respiratory Disease: 1970–2005. *J. Am. Vet. Med. Assoc.* (in press).

Osoro, K., A. Mateos-Sanz, P. Frutos, U. Garcia, L. M. Ortega-Mora, L. M. Ferreira, R. Celaya, and I. Ferre, 2007: Anthelmintic and nutritional effects of heather supplementation on Cashmere goats grazing perennial ryegrass-white clover pastures. *J. Anim. Sci.* 85, 861–870.

Ottenbacher, K. J., H. R. Ottenbacher, L. Tooth, and G. V. Ostir, 2004: A review of two journals found that articles using multivariable logistic regression frequently did not report commonly recommended assumptions. *J. Clin. Epidemiol.* 57, 1147–1152.

Oxman, A. D., and G. H. Guyatt, 1992: A consumer's guide to subgroup analyses. *Ann. Intern. Med.* 116, 78–84.

Parker, K. I., C. W. Compton, F. M. Anniss, C. Heuer, and S. McDougall, 2008: Quarter-level analysis of subclinical and clinical mastitis in primiparous heifers following the use of a teat sealant or an injectable antibiotic, or both, precalving. *J. Dairy Sci.* 91, 169–181.

Pithua, P., S. J. Wells, S. M. Godden, and E. A. Raizman, 2009: Clinical trial on type of calving pen and the risk of disease in Holstein calves during the first 90d of life. *Prev. Vet. Med.* 89, 8–15.

Plint, A. C., D. Moher, A. Morrison, K. Schulz, D. G. Altman, C. Hill, and I. Gaboury, 2006: Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med. J. Aust.* 185, 263–267.

Pocock, S. J., M. D. Hughes, and R. J. Lee, 1987: Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N. Engl. J. Med.* 317, 426–432.

Ribble, C. S., 1990: Assessing vaccine efficacy. *Can. Vet. J.* 31, 679–681.

Rothwell, P. M., 2005: External validity of randomised controlled trials: ''to whom do the results of this trial apply?''. *Lancet* 365, 82–93.

Roy, J. P., D. Harvey, A. M. Belanger, and S. Buczinski, 2008: Comparison of 2-step laparoscopy-guided abomasopexy versus omentopexy via right flank laparotomy for the treatment of dairy cows with left displacement of the abomasum in on-farm settings. *J. Am. Vet. Med. Assoc.* 232, 1700–1706.

Sanford, C. J., G. P. Keefe, I. R. Dohoo, K. E. Leslie, R. T. Dingwell, L. DesCoteaux, and H. W. Barkema, 2006: Efficacy of using an internal teat sealer to prevent new intramammary infections in nonlactating dairy cattle. *J. Am. Vet. Med. Assoc.* 228, 1565–1573.

Sargeant, J. M., A. Rajic, S. Read, and A. Ohlsson, 2006: The process of systematic review and its application in agri-food public-health. *Prev. Vet. Med.* 75, 141–151.

Sargeant, J. M., R. Elgie, J. Valcour, J. Saint-Onge, A. Thompson, P. Marcynuk, and K. Snedeker, 2009a: Methodological quality and completeness of reporting in clinical trials conducted in livestock species. *Prev. Vet. Med.* 91, 107–115.

Sargeant, J. M., J. Saint-Onge, J. Valcour, A. Thompson, R. Elgie, K. Snedeker, and P. Marcynuk, 2009b: Quality of

reporting in clinical trials of pre-harvest food safety interventions and associations with treatment effect. *Foodbourne Pathog. Dis.* 6, 989–999.

Schulz, K. F., and D. A. Grimes, 2002: Generation of allocation sequences in randomised trials: chance, not choice. *Lancet* 359, 515–519.

Schulz, K. F., and D. A. Grimes, 2005: Multiplicity in randomised trials II: subgroup and interim analyses. *Lancet* 365, 1657–1661.

Schulz, K. F., I. Chalmers, D. A. Grimes, and D. G. Altman, 1994: Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 272, 125–128.

Schulz, K. F., I. Chalmers, R. J. Hayes, and D. G. Altman, 1995: Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 273, 408–412.

Schunicht, O. C., C. W. Booker, P. T. Guichon, G. K. Jim, B. K. Wildman, B. W. Hill, T. I. Ward, and S. W. Bauck, 2002: An evaluation of the relative efficacy of a new formulation of oxytetracycline for the treatment of undifferentiated fever in feedlot calves in western Canada. *Can. Vet. J.* 43, 940–945.

Seeger, T., H. Kumper, K. Failing, and K. Doll, 2006: Comparison of laparoscopic-guided abomasopexy versus omentopexy via right flank laparotomy for the treatment of left abomasal displacement in dairy cows. *Am. J. Vet. Res.* 67, 472–478.

Senn, S., 1995: Base Logic: tests of baseline balance in randomized clinical trials. *Clin. Res. Regul. Aff.* 12, 171–182.

Sithole, F., I. Dohoo, K. Leslie, L. DesCoteaux, S. Godden, J. Campbell, G. Keefe, and J. Sanchez, 2006: Effect of eprinomectin pour-on treatment around calving on reproduction parameters in adult dairy cows with limited outdoor exposure. *Prev. Vet. Med.* 75, 267–279.

Smith, D. R., R. A. Moxley, R. E. Peterson, T. J. Klopfenstein, G. E. Erickson, G. Bretschneider, E. M. Berberov, and S. Clowser, 2009: A two-dose regimen of a vaccine against type III secreted proteins reduced Escherichia coli O157:H7 colonization of the terminal rectum in beef cattle in commercial feedlots. *Foodborne Pathog. Dis* 6, 155–161.

Sonis, J., and J. Joines, 1994: The quality of clinical trials published in The Journal of Family Practice, 1974-1991. *J. Fam. Pract.* 39, 225–235.

St-Pierre, N. R., 2007: Design and analysis of pen studies in the animal sciences. *J. Dairy Sci.* 90(Suppl 1), E87–E99.

Taveros, A. A., and S. J. More, 2001: A field trial of the effect of improved piglet management on smallholder sow productivity in the Philippines. *Prev. Vet. Med.* 49, 235–247.

Tetrault, J. M., M. Sauler, C. K. Wells, and J. Concato, 2008: Reporting of multivariable methods in the medical literature. *J. Investig. Med.* 56, 954–957.

Thompson Reuters, 2009: Clinical and Laboratory Standards Institute. http://www. techstreet. com/info/clsi. tmpl (last accessed on 6 November 2009).

Treasure, T., and K. D. MacRae, 1998: Minimisation: the platinum standard for trials? Randomisation doesn't guarantee similarity of groups; minimisation does. *BMJ* 317, 362–363.

Tukey, J. W., 1977: Some thoughts on clinical trials, especially problems of multiplicity. *Science* 198, 679–684.

VanDonkersgoed, J., J. Merrill, and S. Hendrick, 2008: Comparative efficacy of tilmicoson versus tulathromycin as a metaphylactic antimicrobial in feedlot calves at moderate risk for respiratory disease. *Vet. Ther.* 9, 291–297.

VanDonkersgoed, J., D. Hancock, D. Rogan, and A. A. Potter, 2005: Escherichia coli O157:H7 vaccine field trial in 9 feedlots in Alberta and Saskatchewan. *Can. Vet. J.* 46, 724–728.

Villarreal, I., D. Maes, T. Meyns, F. Gebruers, D. Calus, F. Pasmans, and F. Haesebrouck, 2009: Infection with a low virulent *Mycoplasma hyopneumoniae* isolate does not protect piglets against subsequent infection with a highly virulent *M. hyopneumoniae* isolate. *Vaccine* 27, 1875–1879.

Wang, R., S. W. Lagakos, J. H. Ware, D. J. Hunter, and J. M. Drazen, 2007: Statistics in medicine – reporting of subgroup analyses in clinical trials. *N. Engl. J. Med.* 357, 2189–2194.

Wellman, N. G., and A. M. O'Connor, 2007: Meta-analysis of treatment of cattle with bovine respiratory disease with tulathromycin. *J. Vet. Pharmacol. Ther.* 30, 234–241.

Wittek, T., L. F. Locher, A. Alkaassem, and P. D. Constable, 2009: Effect of surgical correction of left displaced abomasum by means of omentopexy via right flank laparotomy or two-step laparoscopy-guided abomasopexy on postoperative abomasal emptying rate in lactating dairy cows. *J. Am. Vet. Med. Assoc.* 234, 652–657.

Yusuf, S., J. Wittes, J. Probstfield, and H. A. Tyroler, 1991: Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 266, 93–98.

Appendix

Participating members of the consensus meeting and steering committee

Robert L. Buchanan (University of Maryland, Center for Food Safety & Security Systems, 0119 Symons Hall, College Park, MD 20742); Cate E. Dewey (Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada); James S. Dickson (Iowa State University, Department of Animal Science, 215F Meat Laboratory, Ames, IA 50011); Ian R. Dohoo (Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, PEI C1A 4P3, Canada); Richard B. Evans (1008 W. Hazelwood Drive, College of Veterinary Medicine/LAC, University of Illinois, Urbana, IL 61802); Brian Fergen (Center for Veterinary Biologics, APHIS, USD,

Ames, IA 50010); Ian A. Gardner (Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, CA 95616); Jeffery T. Gray (Department of Microbiology, 3200 Grand Avenue, Des Moines University, Des Moines, IA 50312); Mattias Greiner (Federal Institute for Risk Assessment, Alt-Marienfelde 17–21, Berlin, Germany D-12277); Greg Keefe (Dairy Health Management, 550 University Avenue, Atlantic Veterinary College University of Prince Edward Island, Charlottetown, PEI C1A 4P3, Canada); Kelly Lechtenberg (Oakland Mercy Hospital, 601 East 2nd Street, Oakland, NE 68045); Sandra L. Lefebvre (American Veterinary Medical Association, 1931 North Meacham Road Suite 100, Schaumburg, IL 60173); Paul S. Morley (Department of Clinical Sciences, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO 80523-1678); Annette M. O'Connor (Veterinary Diagnostic and Production Animal Med, Veterinary Medicine Research Institute Building 4, Iowa State University, Ames, IA 50011); Alex Ramirez (Veterinary Diagnostic and Production Animal Medicine, 2231 Lloyd Veterinary Medicine Center, Iowa State University, Ames, IA 50011); Bradley J. Rauch (VM Quality Milk Production Svc, 22 Thornwood Drive, Cornell University, Ithaca, NY 14853); Susan C. Read (Laboratory for Foodbourne Zoonoses, 110 Stone Road West, Public Health Agency of Canada, Guelph, ON N1G 3W4, Canada); Jan M. Sargeant (Centre for Public Health and Zoonoses, and Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada); William Sischo (Veterinary Clinical Sciences, PO Box 647060, Washington State University, Pullman, WA 99164-7060); David R. Smith (Veterinary and Biomedical Sciences, University of Nebraska-Lincoln, PO Box 830905, Lincoln, NE 68583); Kate Snedeker (Post-doctoral fellow, Centre for Public Health and Zoonoses, Ontario Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada,); John N. Sofos (Colorado State University, Department of Animal Sciences, 1171 Campus Delivery, Fort Collins, CO 80523-1171); Mary E. Torrence (USDA-ARS, GWCC- 4-2194, 5601 Sunnyside Avenue, Beltsville, MD 20870); Michael P. Ward (Faculty of Veterinary Science, The University of Sydney, NSW, Australia 2006); Robert W. Wills (Department of Pathobiology and Population Medicine, 240 Wise Center Drive, PO Box 6100, Mississippi State University College of Veterinary Medicine, MS 39762-6100) and an unnamed government official.