Research in Veterinary Science 186 (2025) 105569

Contents lists available at ScienceDirect

Research in Veterinary Science

journal homepage: www.elsevier.com/locate/rvsc

How many subjects are enough in a veterinary trial?—Literature review and insights from industrial statisticians

Zikun Wang [a], Qi Cao [b,*], Qingzhi Liu [a,c], Divine Dufe [b], Pieter Wouters [b], Lijuan Deng [d], Annpey Pong [a]

[a] Department of Global Statistics and Data Management, Merck Animal Health, 126 E. Lincoln Ave., P.O. Box 2000, Rahway, NJ 07065, USA
[b] Department of Global Statistics and Data Management, MSD Animal Health, P.O. Box 31, 5830 AA Boxmeer, the Netherlands
[c] Department of Biostatistics, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109, USA
[d] Department of Global Statistics and Data Management, Merck Animal Health, 21401 West Center Road, Elkhorn, NE 68022, USA

ARTICLE INFO

Keywords:
Sample size determination
Power calculation
Veterinary clinical trials
Veterinary medicine

ABSTRACT

To approve veterinary medicines, regulatory agencies require substantial evidence of safety and effectiveness from well-designed trials. A critical consideration in planning a veterinary trial is to determine the requisite number of subjects needed to achieve the trial objective(s). Sample sizes for preclinical studies usually follow specific guidelines, whereas clinical studies should follow more rigorous requirements to determine the optimal sample size for detecting significant and meaningful treatment effects. This paper presents a descriptive literature review aiming to explore current reported practices in sample size calculation for published veterinary clinical trials. The review included articles published in 12 top-ranking veterinary health journals from January 1, 2013, to June 4, 2023, sourced from PubMed and Medline regarding animal trials. Initially, 294 articles were identified, of which 98 (33 %) focused on veterinary clinical studies. Among these 98 papers, only 24 (24 %) provided detailed methods on sample size calculations, with individual animals as experimental units across at least two arms. The predominant animal species studied in these trials were canines (39 %), cattle (32 %), and swine (21 %), with most studies aiming for a statistical power of at least 80 %. Unlike human clinical trials, the paper found that statistically rigorous sample size calculations were less commonly reported in animal clinical trials. Our paper provides recommendations for veterinary clinical trial practitioners and offers insights into how sample size determination can be properly conducted and reported. Furthermore, this article extends to discuss practical issues in sample size determination for preclinical studies.

1. Introduction

The demand for evidence-based medicine is more crucial nowadays than ever for improving both human and animal health. Under these circumstances, veterinary trials conducted by pharmaceutical companies provide data-driven evidence to support new animal drug product claims submitted to regulatory authorities, eventually enhancing animal health and welfare while boosting livestock animals' productivity.

As animal drug development progresses through various stages, each phase necessitates different levels of evidence, resulting in variations in the amount of data needed (Singh et al., 2023). Compared to human preclinical trials, which are conducted in laboratory animals before progressing to testing in human trials, preclinical trials for animal drug development are often conducted in controlled lab environments using the target species directly, with the same species continuing into the later-conducted clinical studies. These trials typically use smaller numbers of animals and aim to select the most promising compounds for clinical testing based on safety and efficacy. At the clinical stage, studies are set out to generate evidence in the field which provides real world evidence with fewer experimental restrictions. Veterinary clinical studies normally fall into two categories: pilot clinical studies, which are exploratory and involve fewer animals to establish proof of concept (e.g., optimal drug dose), and pivotal clinical studies, which are confirmatory and require substantially more data to validate lab findings and confirm the safety and efficacy of the optimal dose identified in the pilot study (Committee for Medicinal Products for Veterinary use, 2021).

During the planning stage of a study, one important question that

* Corresponding author.
E-mail address: qi.cao2@merck.com (Q. Cao).

https://doi.org/10.1016/j.rvsc.2025.105569
Received 22 November 2024; Received in revised form 28 January 2025; Accepted 9 February 2025
Available online 10 February 2025

arises is, "How many subjects are needed?" To address this question, many factors must be considered, such as the study objective, study design, primary endpoint, regulatory requirements, clinically meaningful difference, and available budget for conducting the trial (Gupta et al., 2016; Kirby et al., 2002). Therefore, determining a sufficient study sample size for providing reliable answers to the biological or clinical questions under investigation becomes critical.

The current practices for determining the sample size without conducting a statistical power calculation are generally referred to in the following two main aspects. Firstly, to protect the rights and wellbeing of research animals, the licensing of animal experiments typically requires an ethical evaluation process for approval (Ghasemi and Dehpour, 2009). This consideration applies to both preclinical animal studies for human health and veterinary studies for animal health. Specific ethical considerations, such as the well-known 3R principles (Replacement, Reduction and Refinement) (Flecknell, 2002)), may influence the number of experimental animals used in study and therefore impact the necessity of statistical justifications for sample size determination. From a scientific perspective, a sufficient sample size to provide a desired power for detecting a clinical meaningful difference would be desirable, but from an animal welfare standpoint, a low number of animals is preferred, which may not be optimal to detect a meaningful difference between treatment groups.

Secondly, different regulatory authorities (i.e., The Food and Drug Administration's Center for Veterinary Medicine (FDA CVM), the United States Department of Agriculture's Center for Veterinary Biologics (USDA CVB), the European Medicines Agency's Committee for Veterinary Medicinal Products (EMA CVMP)), and scientific associations (i.e., World Association for the Advancement of Veterinary Parasitology (WAAVP)) have published guidelines emphasizing the importance of sample size determination to ensure robust study design and reliable study result (Burden et al., 2024, Committee for Medicinal Products for Veterinary use, 2021, Otranto et al., 2021, Shere, 2014). Some documents provided specific recommendations for sample size numbers based on disease and animal species. For instance, for animal vaccination studies in Europe, different monographs in European Pharmacopoeia (Ph. Eur.) (European Directorate for the Quality of Medicines and HealthCare, 2016) are used to determine the number of animals included in laboratory studies based on the active components in the product. For pharmaceutical studies, the WAAVP provides guidelines on the minimum number of animals required for each group when designing different types of studies (Burden et al., 2024; Geurden et al., 2022; Otranto et al., 2021).

Although sample size determination based on current guidelines, which often lack power calculations, still plays an important role in animal trials, it is crucial to realize that the sample size recommended only by these guidelines may not be optimal in terms of evaluating the true effectiveness of study medicine. Note that study design varies based on the study objectives, while the numbers of animals suggested in the guidelines are generally given within certain ranges which do not take any other critical factors into considerations, include but not limited to variability of primary endpoint and availability of new efficacy endpoints due to improved research over the years.

An inadequate sample size can result in a failed trial despite treatment effectiveness, while an excessively large sample size can lead to unnecessary waste of time, money, and unethically inclusion of experimental animals. In the current animal health study literature, rigorous sample size determination based on statistical justifications and description of these methods does not receive sufficient attention. Determining the sample size with a statistical power calculation (referred to as sample size calculation hereafter) ensures that the study has adequate power to detect meaningful clinical effects or differences between treatment groups, minimizing the risk of drawing erroneous study conclusions due to insufficient or excessive sample sizes. In addition, determining the appropriate sample size can optimize the usage of resources and reduce unnecessary animal use.

A review article concluded that there was a very limited veterinary clinical trials clearly defined the primary outcome and properly considered sample size calculation (Wareham et al., 2017). The conclusion was made based on the articles that published in the year 2011 and focused more on quality of veterinary clinical trials with sample size determination as one of the quality standards. In this paper, we aimed to explore current practices in sample size calculation for veterinary clinical trials through a more recent literature review over ten years publication window from 2013 to 2023 which offers an up-to-date assessment of practice and highlight the important elements shown in the literature. Additionally, we provided recommendations for veterinary clinical trial practitioners and offer insights of how sample size determination can be properly conducted. Furthermore, the practical issues in sample size determination for preclinical studies were discussed.

2. Methods

A descriptive literature review (Paré et al., 2015) was used to select relevant papers describing sample size calculations in veterinary clinical trials. Our search strategy aimed to yield a robust collection of articles for review, thereby providing a comprehensive and contemporary understanding of current practice in the animal health field.

The following predefined criteria were established for inclusion in our descriptive review:

- The study must involve veterinary clinical trials with at least two groups
- The study should focus on individual animal as the experimental unit
- The study should report detailed information (i.e., calculation methods, effect size, adequately power ($\geq$ 80 %) (Serdar et al., 2021)) on sample size calculation
- The study must be published in one of the top 12 vet health journals as specified below
- The publication date of the study should fall between January 1, 2013, and June 4, 2023
- Review papers are to be excluded from consideration

The databases of PubMed and Medline were used in our review. To ensure the inclusion of top-ranking papers, our focus was narrowed to 12 specific journals categorized under general veterinary, animal science and zoology based on their 5-year impact factor and CiteScore according to Scopus (Goodman and Deis, 2005). The following journals were included: Parasites & Vectors, Veterinary Research, Preventive Veterinary Medicine, Veterinary Pathology, Journal of Veterinary Internal Medicine, Zoonoses and Public Health, Veterinary Parasitology, BMC Veterinary Research, Journal of the American Veterinary Medical Association, The Veterinary Journal, Veterinary Record Open, and Veterinary Dermatology. Search terms incorporated a broad range of keywords among the above included journals, including 'animal study', 'animal trial', 'animal experiment', 'field study', and 'field trial' which were employed in the title and abstract fields and 'individual animal' in the full text of these databases. Additional filters were applied to narrow the search to articles published in English, full-text articles, books and documents, and specific types of studies (e.g., Clinical Trial, Comparative Study, Clinical Trial, Veterinary, Controlled Clinical Trial, Multicenter Study, Randomized Controlled Trial, Validation Study). The search was further delimited to articles published between January 1, 2013, and June 4, 2023.

Two independent reviewers (Z.W., Q.C.) reviewed each article especially on title, abstract, introduction, and methods sections to see whether the inclusion criteria were met and whether any sample size calculation method was specified. If disagreement existed between the two reviewers, a third reviewer (P.W.) was consulted. When necessary, articles can be further excluded after reviewing the relevant full text. The reason for exclusion was clearly described in a flowchart that

outlines the entire review process (Fig. 1).

3. Results

3.1. Article selection and screening process

The initial search across PubMed and Medline databases, based on the selected journals, keywords, and the specified time period, yielded 294 publications on animal trials. After title and abstract screening, the majority of records (185) were excluded as they were unrelated to veterinary clinical trials, such as preclinical animal studies conducted for human health purposes. Additionally, two articles were unavailable in full-text format, and 9 were review papers. All these above articles were excluded. This process narrowed the list down to 98 published articles focused on animal clinical trials, which were then reviewed in detail. Of these 98 articles, 68 were subsequently excluded among which 36 articles (53 %) did not report the statistical method used to calculate sample size, 19 articles (28 %) only mentioned the sample size was chosen based on guideline without providing further detail, and 13 articles (19 %) did not meet the inclusion criteria of our review after further screening, leaving 30 full-text articles that clearly reported sample size calculation method to be further assessed. Six of these articles failed to meet the inclusion criteria and were therefore excluded from the final selected studies. The reasons for exclusion of these six articles were as follows: one was an observational study, one was a questionnaire-based study, one was a clinical trial with only a single treatment arm, one study had an inadequate power (<80 %), and two studies did not use individual animals as the experimental units. In total, there are 24 studies that met all the criteria for inclusion and were incorporated into the extraction Table 1 with relevant detailed information.

3.2. Characteristics of included studies

3.2.1. Overview of included studies

Among these 24 studies, 29 % were published in Veterinary Record Open, followed by Journal of Veterinary Internal Medicine (17 %), Parasites & Vectors (17 %), BMC Veterinary Research (12 %), and 25 % from 8 other journals. 11 studies were published between 2013 and 2017 and 13 studies were published between 2018 and the first half of 2023. A slight increase in the reporting sample size calculation has been seen in the past 10 years. Geographically, most of the studies were conducted in Europe (15 studies), with Belgium (5 studies) and Finland (3 studies) being the most frequent countries. Six studies took place in North America, with five of them conducted in the USA. The study species were 46 % ($N = 11$) on dogs, 29 % ($N = 7$) on cattle, 17 % ($N = 4$) on swine, and 4 % (N = 1) each on horses and lambs.

All articles utilized a parallel study design, where subjects took one treatment and remained in their assigned treatment groups after randomization till the end of the study. Moreover, most of the studies ($N = 17$) were designated for evaluating treatment efficacy. The remaining studies (N = 7) were designated for evaluating both efficacy and safety. Superiority trials constituted 92 % ($N = 22$) of the 24 included studies, while non-inferiority trials made up 8 % (N = 2). The most frequently used software for calculating sample size was the SAS® system (Institute, Cary, North Carolina, USA) ($N = 3$).

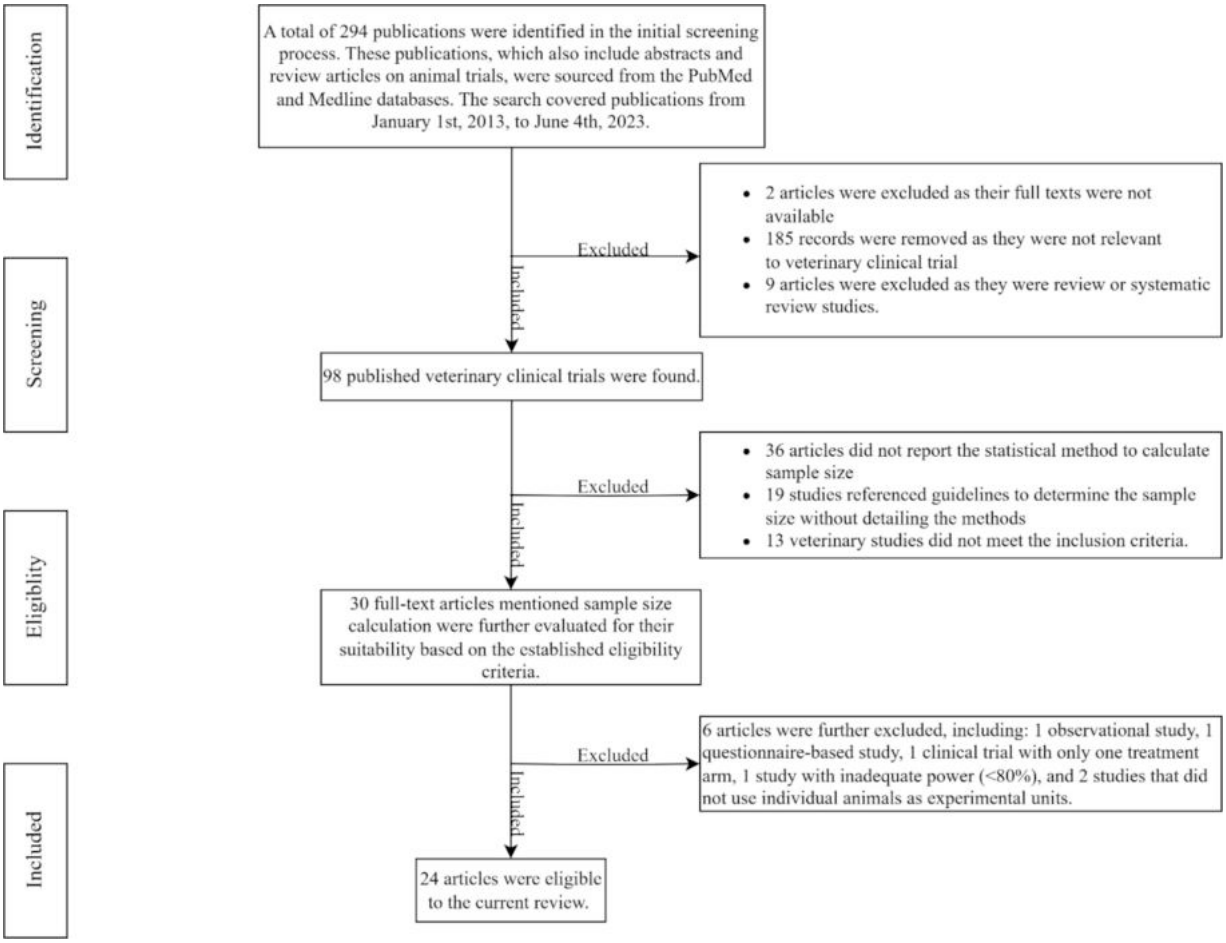

Fig. 1. Flowchart of the study selection process.

Table 1
Summary of Eligible Articles and Components of Sample Size Calculations in Animal Clinical Trials.

	Authors and Year	Journal	Country	Species	Aim	Clinical Study Design (efficacy/ safety)	Clinical Study Hypothesis	Group Details	Sample Size Determination Method and Parameters (Distribution)	Primary Parameter and (Inferential Model)	Software Used
1	Almawly et al., 2013	Veterinary Parasitol	New Zealand	Cattle/ Calves	The anti-Cryptosporidium preventive efficacy of the compound in calves on a farm enzootically infected with *C. parvum*, bovine rotavirus and Salmonella Typhimurium.	Randomized controlled field trial (efficacy trial)	Superiority	Group 1: Full dose treated with 8 mL Halocur); Group 2: Half dose treated with 4 mL Halocur); Group 3: Placebo-control;	Log10-transformed mean number of oocysts per gram of feces of 4.3 for the full dose and 6.2 for the untreated-control group at the peak of shedding, and a common standard deviation (SD) of 1.2 (log-normal distribution); Group sizes of 8 calves per group achieved 84 % power to reject the null hypothesis that both group means are 6.2.	Not clearly mentioned	PASS software (NCSS, Kaysville, UT)
2	Arsenakis et al., 2017	Veterinary Re cord Open	Belgium	Swine/ Sows	Efficacy of *Mycoplasma hyopneumoniae* vaccination in pig herd with mixed respiratory disease	Randomized efficacy trial (efficacy trial)	Superiority	V1 + V2: a single-dose injection of a commercial Mhyopneumoniae bacterin vaccine (Ingelvac MycoFLEX, Boehringer Ingelheim); V1: Vaccinated before weaning); V2: Vaccinated at the day of weaning); NV: Non-vaccinated group.	276 per group for a difference of 19 g (SD =80) in average daily weight gain (ADG) (normal distribution); 3.2 points (SD = 13) in lung lesion score (normal distribution) with 95 % certainty and 80 % power.	ADG (ANOVA) and lung lesion score (non-parametric Kruskal-Wallis ANOVA).	Win Episcope 2.0, CLIVE, Edinburgh, UK
3	Arsenakis et al., 2018	Veterinary Record Open	Belgium	Swine/ Sows	Efficacy of vaccinating gestating sows facing recurrent outbreaks of exudative epidermitis	Randomized efficacy trial (efficacy trial)	Superiority	Two batches were vaccinated (V) against S hyicus and two remained non-vaccinated (NV).	1800 pigs per farrowing batch with 95 % certainty and 80 % power to detect a difference of 1.7 % points (binomial distribution).	Morbidity and Mortality (Logistic regression)	IBM SPSS Sample Power V.3, Illinois, USA
4	Chatzis et al., 2017	Parasites & Vectors	Greece	Canine/ Dogs	Efficacy of a fixed combination of permethrin and fipronil for the treatment and prevention of flea infestation	Randomized, blinded, placebo-controlled trial (efficacy trial)	Superiority	Group A: Spot-on solution containing 54.5 % permethrin and 6.1 % fipronil (Effitix®), at the label dose; Group B: Sham-treated with empty Effitix® pipettes.	The study power was 80 % to detect a 54 % difference between the two groups in the number of dogs with zero flea counts (binomial); 80 % power to detect a 21 % difference of flea counts between the two groups, assuming a SD of 0.2, α = 5 % (normal distribution).	The percentage of dogs with zero flea counts (Fisher's exact test)	Not clearly mentioned
5	Dantas-Torres et al., 2013	Parasites & Vectors	Italy	Canine/ Young dogs	Efficacy of an imidacloprid/ flumethrin collar against fleas, ticks and tick-borne pathogens	A parallel group-designed, randomized, controlled efficacy (efficacy trial)	Superiority	Groups A: Dogs with collars containing combination of imidacloprid 10 % and flumethrin 4.5 %; Group B: Untreated control dogs.	49 per group. Expected Incidence in Group A: 5 % Expected Incidence in Group B: 15 % Power of the Study: 85 % Level of Confidence (or significance level): 95 % (binomial distribution).	Not clearly mentioned	WinEpi (http://www.winepi.net/uk/index.htm)

(*continued on next page*)

Table 1 (*continued*)

	Authors and Year	Journal	Country	Species	Aim	Clinical Study Design (efficacy/ safety)	Clinical Study Hypothesis	Group Details	Sample Size Determination Method and Parameters (Distribution)	Primary Parameter and (Inferential Model)	Software Used
6	Del Pozo Sacristan et al., 2014	Veterinary Re cord Open	Belgium	Swine/ Piglets	Efficacy of early *Mycoplasma hyopneumoniae* vaccination	Field efficacy (efficacy trial)	Superiority	V1: Vaccinated at seven days of age with a commercial inactivated M hyopneumoniae vaccine (Stellamune Once, Elanco Animal Health); V2: Vaccinated at 21 days of age using the same vaccine as V1; NV: Non-vaccinated.	180 pigs/group allows a difference of 20 g in Average Daily Gain (ADG) with a SD of 70 g. To achieve this difference with a 95 % certainty and 80 % statistical power (normal distribution).	Not clearly mentioned	Win Episcope 2.0, CLIVE, Edinburgh, UK
7	Forster et al., 2018	BMC Veterinary Research	USA	Canine/ Dogs	Efficacy and safety of a terbinafine, florfenicol and betamethasone topical ear formulation for the treatment of bacterial and/or fungal otitis externa	Randomized double-blinded placebo (vehicle)-controlled multicenter parallel trial; (efficacy and safety trial)	Superiority	Treatment: The topical application of 1 mL of the active product (Osurnia); Control: Vehicle of treatment (placebo control), once on Day 0 and a second time on Day 7 (±2 days).	Anticipated recruitment was 225 dogs (minimum 150 evaluable cases) for inclusion in a 2:1 ratio of active-versus-vehicle. To achieve approximately 85 % power in identifying a treatment success rate of 70 % with Osurnia and 45 % with the placebo (binomial distribution).	Rate of treatment success (RTS) on Day 45 (Generalized Linear Mixed Model with Logit link function)	Not clearly mentioned
8	Hanks et al., 2021	Veterinary Parasitol	Australia	Lambs	Effect of pasture molluscicide treatment on the prevalence and severity of small lungworm infection and the productivity of lambs grazing improved pastures	Randomized control field trial (efficacy trial)	Superiority	Treatment: According to the manufacturer's instructions for heavily snail-infested pasture; Control: Bait was not applied to the vast majority of control paddocks.	130 per group enabled detection of a difference in carcass weight of 700 g between experimental groups assuming a variance of 2.89 kg^2, power of 80 %, and desired confidence of 95 % (normal distribution).	Not clearly mentioned	Not clearly mentioned
9	Jourquin et al., 2022	Journal of Vet Intern Medicine	Belgium	Cattle/ Calves	Efficacy of florfenicol and oxytetracycline in calf pneumonia	A randomized field trial (efficacy trial)	Non-inferiority	Group 1: Treated for *M. bovis* pneumonia with FF (Selectan, Hipra; Group 2: Treated for *M. bovis* pneumonia with with OTC (Engemycine, MSD).	130 animals in total allowed detection of a 10 % difference in cure rate between the 2 groups, with 80 % power ($\alpha = 5$ %) and a non-inferiority margin of 10 % (binomial distribution); Using an expected SD in healing time of 4 days allowed for detection of differences in healing time of 2 days between FF and OTC (normal distribution.	Cure rate (Multivariable Logistic Regression with cure as binary outcome) and healing time (Cox proportional hazards model)	Not clearly mentioned
10	Kneipp et al., 2023	Preventive Veterinary Medicine	Australia	Cattle/ Beef calves	Efficacy of a commercial vaccine for pinkeye	A randomized control trial (efficacy trial)	Superiority	Treatment: 2 mL pinkeye vaccine (Piliguard® (Intervet Australia Pty Ltd. known as MSD Animal Health)); Control: No treatment.	157 animals per group allowed 5 % of the vaccinated and 15 % of the control animals may acquire the disease with 80 % power and with 5 % level of significance (binomial distribution).	The incidence of naturally occurring pinkeye (Generalized Linear Mixed Models)	Not clearly mentioned

(*continued on next page*)

Table 1 (*continued*)

	Authors and Year	Journal	Country	Species	Aim	Clinical Study Design (efficacy/ safety)	Clinical Study Hypothesis	Group Details	Sample Size Determination Method and Parameters (Distribution)	Primary Parameter and (Inferential Model)	Software Used
11	Korpivaara et al., 2017	Veterinary Re ecord Open	Germany and Finland	Canine/ Dogs	Efficacy of dexmedetomidine in alleviation of acute anxiety and fear associated with noise	A randomized, double-blind, placebo-controlled, parallel-group, clinical-field study (efficacy and safety trial)	Superiority	Group 1: 0.1 mg/mL dexmedetomidine oromucosal gel at a dose of 125 mg/m^2; Group 2: An equivalent volume of placebo gel.	To detect a difference in the five categories of the first co-primary variable, 65 dogs in each group allowed 95 % power (α = 5 %, 2-sides), dropout rate of 10 %, 70 dogs were to be recruited for each group (multinomial distribution).	The overall effect of study treatment (multinomial score) (Generalized Linear Model with cumulative logit as a link function)	Sample size calculation was based on the previous study (Korpivaara and others 2014)
12	Korpivaara et al., 2021	Veterinary Re cord Open	Finland	Canine/ Dogs	Clinical safety and efficacy of dexmedetomidine for alleviation of acute anxiety manifested as fear or fear induced fractiousness	A double-blinded, randomized, placebo-controlled, parallel-group, multicenter, clinical field study (efficacy and safety trial)	Superiority	Group 1: Dexmedetomidine 0.1 mg/g Oromucosal gel at a dose of 125 μg/m^2; Group 2: Group 1 at a dose of 250 μg/m^2; Group 3: Placebo gel.	A sample size of 72 dogs was calculated to provide 94 % power to detect a difference between the combined dexmedetomidine groups and placebo at α = 5 % level, using a one-sided test (multinomial distribution).	The investigator's ability to perform the intended procedure (Generalized Linear Model using a cumulative logit link function.)	Not clearly mentioned
13	Larsen et al., 2023	Preventive Veterinary Medicine	Denmark	Swine/ Piglets	Efficacy of autogenous vaccines and iodine application on the risk of Umbilical Outpouchings (UO) development	A blinded, randomized clinical field trial (efficacy trial)	Superiority	PO: Piglets from placebo-vaccinated sows, no umbilical treatment; P1: Piglets from placebo-vaccinated sows, with iodine umbilical treatment; VO: Piglets from autogenously vaccinated sows, no umbilical treatment; VI: Piglets from autogenously vaccinated sows, with iodine umbilical treatment.	Presumed reduction in umbilical outpouchings (UO) prevalence from 5.0 % to 3.5 %, 95 % significance level, 80 % power, one-sided testing resulted in 1125 pigs per group (binomial distribution).	Not clearly mentioned (however only UO was mentioned in statistical section using Logistic Regression)	R version 4.0.4
14	Lascelles et al., 2016	BMC Veterinary Research	USA	Canine/ Dogs	Efficacy of bupivacaine liposome injectable suspension for the provision of post-surgical analgesia	A masked, randomized, placebo-controlled, multi-center pilot field study (efficacy and safety trial)	Superiority	Treatment: A novel bupivacaine liposome injectable suspension (AT-003); Control: Treated with Saline.	The enrollment target was at least 40 evaluable cases (20 AT-003 and 20 placebo). This number allowed to detect a 0.37 difference (pooled SD ±0.30) in the CMPS-SF assessment at 24 h post-administration with 90 % power (α = 5 %, 2-sides) (normal distribution).	The Glasgow Composite Measure Pain Scale (Short Form) (CMPS-SF) (The overall score of CMPS-SF with the secondary efficacy parameter (Surgical site manipulation score (SSMS) was analyzed by repeated measures ANOVA)	Not clearly mentioned
15	Masmeijer et al., 2019	Journal of Vet Intern Medicine	Belgium	Cattle/ Dairy calves	Effects of body weight and short transport on stress and immune variables	A randomized controlled field trial (efficacy trial)	Superiority	Group 1: Underweight/ no transport (LOWCON); Group 2: Underweight/	Sample size was based on one of the primary variables. To detect a difference of 2400 relative	Mean value of the neutrophil Reactive oxygen species (ROS) production test	SAS System® v9.4 (SAS Institute, Cary,

(*continued on next page*)

Table 1 (*continued*)

	Authors and Year	Journal	Country	Species	Aim	Clinical Study Design (efficacy/ safety)	Clinical Study Hypothesis	Group Details	Sample Size Determination Method and Parameters (Distribution)	Primary Parameter and (Inferential Model)	Software Used
								transport (LOWTRANS); Group 3: Normal weight/no transport (NORMCON); Group 4: Normal weight/transport (NORMTRANS).	RLU (SD of 1200 RLU) between 2 groups, with 95 % confidence interval, 80 % study power, 5 animals in each group were needed (normal distribution).	(repeated measurement Linear Mixed Model)	North Carolina, USA)
16	Masset et al., 2020	Veterinary Journal	France	Cattle/ Beef calves	Efficacy of two intranasal vaccines for the control of bovine respiratory disease	A randomized non-inferiority multicenter trial (efficacy trial)	Non-inferiority	Vaccine A: Bovalto Respi Intranasal, Boehringer Ingelheim; Vaccine B: Rispoval RS + PI3 Intranasal, Zoetis.	446 calve per group to detect non-inferiority assuming $\alpha = 5$ %, $\beta = 0.2$ (=power 0.8). δ (non-inferiority margin) = 0.05; and a prevalence of BRD in the active control vaccine group B of 10 % (binomial distribution).	Prevalence of calves treated for BRD during the study period (Mixed Logistic Regression Model with multiple covariates, optimal model generated after variable selection)	R package 'TrialSize'
17	Meadows et al., 2014	Parasites & Vectors	USA	Canine/ Dogs	Efficacy of fluralaner tablets in controlling canine flea infestations	An investigator-blinded, multicentric, positive-controlled study (efficacy and safety trial)	Superiority	Treatment: Fluralaner tablets; Control: Three sequential treatments of orally administered spinosad.	No more than 10 households would be sufficient to provide 80 % power to demonstrate statistically significant (p-value <0.05) reduction of flea counts from baseline levels; The study targeted enrolling 100 households to be allocated to the fluralaner treatment group and 33 households to the control (spinosad + amitraz) group (log-normal distribution).	Live flea counts in primary dogs as the experimental unit at weeks 4, 8, and 12 (Days 28, 56, and 84) compared to baseline counts (log-transformed data were analyzed by a Mixed Linear Model with repeated measures).	Not clearly mentioned)
18	Nixon et al., 2019	Journal of Vet Intern Medicine	UK and Ireland	Canine/ Dogs	Efficacy of an orally administered anti-diarrheal probiotic paste in dogs with acute diarrhea	A double-blinded, placebo-controlled, randomized, multicenter clinical field study (efficacy trial)	Superiority	Treatment: a commercially available antidiarrheal probiotic paste (ADPP) containing *E. faecium* 4b1707 (Pro-Kolin Advanced; PKA); Control: Placebo Group.	43 cases per group to detect a difference in survival curves for dogs being diarrhea-free on day 3, with $\alpha = 5$ % and power of 80 % using a 2-sided log-rank test. The sample size calculation was based on 80 % of dogs in the ADPP group and 50 % of dogs in the Placebo group being free from diarrhea on Day 3. A target of 50 cases per group was set (time-to-event data).	Duration of diarrhea (ITT population using Cox proportional hazard multivariable regression, PP population using the Mann-Whitney test)	SAS (nQuery Advisor Version 7.0)
19	Pithua et al., 2013	American Veterinary Medical Association	USA	Cattle/ Heifer calves	Efficacy of a lacteal-derived colostrum replacer (LDCR) for the prevention of failure of	A randomized field trial (efficacy trial)	Superiority	Treatment: Fed with LDCR; Control: Fed with	474 calves (237 calves per treatment group) to allow incidence rates of failure of passive transfer of	Not clearly mentioned	Stata Corp for Windows (StataCorp LP, Texas, USA)

(*continued on next page*)

Table 1 (*continued*)

	Authors and Year	Journal	Country	Species	Aim	Clinical Study Design (efficacy/ safety)	Clinical Study Hypothesis	Group Details	Sample Size Determination Method and Parameters (Distribution)	Primary Parameter and (Inferential Model)	Software Used
					passive transfer of immunity (FPT)			pooled maternal colostrum (MC).	immunity (FPT) in calves fed LDCR being 10 % and pooled MC being 19 % with $\alpha = 5$ % and ≥ 80 % study power (binomial distribution).		
20	Salichs et al., 2022	Veterinary Re cord Open	Spain and France	Canine/ Dogs	Efficacy and safety of enflicoxib for canine osteoarthritis	A prospective, multisite, blinded, randomized, controlled, parallel-group field study (efficacy and safety trial)	Superiority	Group 1: Oral treated with enflicoxib at 4 mg/ kg; Group 2: Oral treated with enflicoxib at 2 mg/ kg; Group 3: Placebo; Group 4: Oral treated with mavacoxib at 2 mg/kg.	The proportion of responders in the placebo group was assumed to be approximately 30 % and 60 % in each treated group. Thus, a sample size of 42 animals per group would provide a statistical power of 80 % when testing for differences by means of a chi-square test (binomial distribution).	Percentage of clinical sum score (CSS) responders at the end of the study (Day 42) (Logistic Regression).	Not clearly mentioned
21	Sandelin et al., 2020	BMC Veterinary Research	Finland	Cattle/ Calves	Effect of intranasal respiratory vaccine on bovine respiratory disease	A randomized blinded, field trial (efficacy trial)	Superiority	Treatment: Vaccinated with a single-use plastic nasal cannula (Zoetis); Control: Unvaccinated;	Study power was 0.9 to detect a difference of 10 % in treatment rate between study groups with 5 % significance level (binomial distribution).	Not clearly mentioned	Stata/MP 14.1 for Windows (StataCorp LP, Texas, USA)
22	Saridomichelakis et al., 2015	Parasites & Vectors	Greece	Canine	Efficacy of spinosad to treat and prevent flea infestations in shepherd dogs living on sheep farms	A randomized, blinded, placebo-controlled trial (efficacy trial)	Superiority	Group A: Spinosad (Comfortis; Elanco Animal Health), at the dose registered in Europe; Group B: Placebo tablets.	The study power was 80 % to detect a 54 % difference between the two groups in the number of dogs with zero flea counts (binomial); 80 % power to detect a 21 % difference of flea counts between the two groups, assuming a SD of 0.2, $\alpha =$ 5 % (normal distribution).	Not clearly mentioned	Not clearly mentioned
23	Schoster et al., 2015	Journal of Vet Intern Medicine	Canada	Equine/ Neonatal foal	Effect of a Probiotic on Prevention of Diarrhea and Clostridium difficile and *Clostridium perfringens* Shedding in Foals	A randomized, placebo-controlled field trial (efficacy trial)	Superiority	Treatment: Probiotic product; Control: Placebo product contained the starch without the bacterial Cultures.	With the assumptions of a 50 % diarrhea incidence, a sample size of 93 animals was calculated to detect a 20 % difference with a power of 0.8 and $\alpha = 5$ % (binomial distribution).	Not clearly mentioned	Not clearly mentioned
24	Suokko et al., 2020	Veterinary Re cord Open	USA	Canine/ Dogs	Efficacy, safety and usability of ropinirole eye drops to induce vomiting in dogs	A randomized, double-blind, placebo-controlled clinical field study (efficacy and safety trial)	Superiority	Treatment: Ropinirole 30 mg/mL ophthalmic solution in a single-use blow fill sealed ampoule; Control: Placebo ophthalmic solution.	90 dogs in the ropinirole treatment group to achieve approximately in 85 % of 10,000 statistical analyses yielding a 95 % CI above 70 % responder rate. The total planned sample size was 120 dogs, of which 25 % of the dogs received	Responder rate of induction of vomiting within 30 min following the initial treatment administration (Generalized Linear Mixed Model with a binomial random	SAS System® v9.4 (SAS Institute, Cary, North Carolina, USA)

(*continued on next page*)

Table 1 (*continued*)

Authors and Year	Journal	Country	Species	Aim	Clinical Study Design (efficacy/ safety)	Clinical Study Hypothesis	Group Details	Sample Size Determination Method and Parameters (Distribution)	Primary Parameter and (Inferential Model)	Software Used
								placebo and so the study was powered to determine whether the responder rate was significantly greater than 70 %. (3:1 ratio).	variable and a logit link function)	

3.2.2. Main parameters and statistical modeling considerations in sample size calculation for included studies

All 24 included studies aimed to achieve at least 80 % power to detect a meaningful difference between the treatment groups. Most of the studies included chose 5 % as the significance level. In most of these included studies, the sample size was calculated based on the primary efficacy endpoints, whereas few studies did not clearly mention their primary efficacy criteria, making it difficult to determine whether sample size calculation was performed based on the primary efficacy endpoints.

For the 22 superiority studies, the choice of parameter distribution and statistical method for sample size calculation varied depending on the study objective and the primary endpoints of the study. When the primary outcomes involved differences in continuous outcomes between groups, a normal distribution was often assigned and the statistical method for sample size calculation would require the mean and standard deviation (SD) as input parameters besides the desired statistical power (e.g., 80 %) and pre-defined significance level (e.g., 5 %). The examples for studies involving continuous outcomes and mean/SD in sample size calculations were body weight changes (Arsenakis et al., 2017; Del Pozo Sacristan et al., 2014), Pain scores (Lascelles et al., 2016), and healing time (Jourquin et al., 2022). From Table 1, Analysis of Variance (ANOVA) model was the common model used for continuous primary endpoint, which matches the statistical method in sample size calculation. For studies where the distribution of the continuous outcome was skewed, a log transformation of the original data was included, as described by Almawly et al. (Almawly et al., 2013)

When the primary outcomes involved binary outcomes (e.g., responder vs non-responder), binomial distribution was normally assumed and the statistical method for sample size calculation requires response rate from each treatment group as the input parameters besides the power and significance level. Examples for studies (Arsenakis et al., 2018; Chatzis et al., 2017; Dantas-Torres et al., 2013; Forster et al., 2018; Jourquin et al., 2022; Kneipp et al., 2023; Masset et al., 2020; Pithua et al., 2013; Salichs et al., 2022; Sandelin et al., 2020; Schoster et al., 2015; Suokko et al., 2020) that involved binomial distributions and rate evaluations in sample size calculations used treatment success rates or specific incidence rates (e.g., mortality, presence or absence of a condition). These studies typically used logistic regression as the most common model for analyzing binary primary endpoints, which aligns with the statistical methods employed in sample size calculations. Besides, multinomial distribution followed by generalized linear model with cumulative logit link and survival analysis were employed for categorical and time-to-event data serving as the primary efficacy criterion, these can be seen in studies by (Korpivaara et al., 2017; Korpivaara et al., 2021; Nixon et al., 2019). A detailed summary is provided in Table 1.

For the two non-inferiority studies (Jourquin et al., 2022; Masset et al., 2020), apart from the power and significance level and other relevant parameters depending on their distributions, a non-inferiority margin was provided to determine the sample size to conclude equivalence between treatment groups within a pre-defined interval.

4. Discussion

4.1. Current practice of sample size calculation in veterinary clinical trials

In the descriptive literature review, we summarized the methods of sample size calculations from 24 veterinary clinical studies published in leading veterinary journals globally over the past decade. An increasing trend to include sample size calculation method in veterinary clinical trials was found with 11 papers published in the first five years and 13 in the last 5.5 years. Although only covered a half year, 2 included articles were published in 2023, which equals the number of included publications in the whole 2018, 2019, 2021 and 2022. Most of the eligible studies in this review focused on dogs, cattle, and swine, with fewer

studies on horses and lambs. There was no study for aquaculture and poultry. This is most likely due to the inclusion criteria in this review that experimental unit is each individual animal, whereas poultry and fish trials typically use groups (e.g., pen, cage) as experimental units. The most frequently used software for sample size calculations among these studies was SAS System® (Institute, Cary, North Carolina, USA). Additionally, our summarized results reveal that most of the eligible studies were designed for evaluating the superiority of the test group, with all these studies aiming for at least 80 % power to detect a statistically significant difference between treatment groups.

The current review found 98 articles related to veterinary clinical trials during the screening process. Out of the 98 articles, 74 were excluded from the final review. The primary reason for exclusion was more than half of the articles failed to report the statistical method to calculate sample size. Particularly, the majority of these excluded studies were conducted in 2018 and focused on dogs, indicating that more established guidelines, such as (Geurden et al., 2022), may exist for canine research compared to other species. This finding also illustrates that sample size calculation in veterinary trials has not yet been as thoroughly described or widely used as expected. Instead, practitioners often rely on numbers from guidelines or previous studies without taking statistical considerations into account. It is promising that in the current review, some papers further optimize their sample size calculations after consideration of practical factors such as drop-out rates. However, a few included articles did not elaborate on the distribution of the primary efficacy outcome when choosing their sample size calculation method or even did not clearly mention what was their primary efficacy criteria, even though they performed the sample size calculation in their studies. This makes it difficult to judge whether all included articles perform sample size calculation based on the primary efficacy criteria.

4.2. Practical recommendations for sample size determination in veterinary trials

Based on the current findings, sample size calculation with thorough descriptions of methods used are currently not widely reported in veterinary literature. The reasons are summarized below. First, the aims of veterinary pre-clinical or pilot clinical studies are often exploratory, which focus on selecting the most promising treatments. When designing veterinary clinical pilot studies, where there is no prior knowledge for the sample size determination, Julious (Julious, 2005) advocated a size of 12 per group as a rule of the thumb. This recommendation is based on feasibility; precision about the mean and variance; and regulatory considerations. Although this publication was targeted to veterinary trials, the statistical principles do generally apply. In relevant situations, the determination of sample size in pre-clinical studies should also consider different country policies about animal welfare (Federal Law Gazette I, p, 1998).

Second, the restrictions of the 3R principles and evidence from well-established guidelines/pharmacopoeia can be considered which limits the statistical justifications on sample size determination. In general, the 3R principles serve as a good harmonization of science and ethics in veterinary field for early phase studies (Ko and Lim, 2021). However, such a low number of animals may not be optimal to detect a meaningful difference between treatment groups.

Last but not least, when well-established guidelines are available, they should also be considered. For instance, the latest WAAVP recommends a minimum of 6–10 animals per group for pilot clinical studies to ensure successful infection of controlled animals and efficacy of treated animals when evaluating the efficacy of parasiticides in reducing the risk of vector-borne pathogen transmission in dogs, cats, and cattle (Burden et al., 2024; Geurden et al., 2022; Otranto et al., 2021).

Although sample size determination based on power calculations is not widely applied or well-documented in pilot clinical trials, it becomes more often used at the confirmatory stage. At confirmatory stage, the safety and efficacy of the active product are more certain, and the relevance of using sufficient study animals to fully prove efficacy is more apparent than earlier phases. As a result, the traditional sample size calculations, even if they require a larger number of animals, are more easily convinced and justified. In practice, factors such as time, budget, and resource availability can limit the maximum sample size achievable, leading to a trade-off between cost-effectiveness and statistical power (Chow et al., 2017). Therefore, researchers/practitioners must balance the ideal sample size with what is feasible within their constraints. On the other hand, accounting for potential dropout rates by increasing sample sizes accordingly ensures that the study retains sufficient power even if some participants withdraw, thereby safeguarding the integrity of the study's results. Hence, conducting sample size calculations is necessary and important as it optimizes the sample size for detecting significant and meaningful treatment effects.

4.3. Theoretical and software recommendations for sample size calculation

From a statistical perspective, several factors affect sample size calculation, including but not limited to the distribution of the primary outcome, study aims, trial types, chosen confidence level (1 - Type I error rate) or desired power (1 - Type II error rate), expected effect size, equivalence/non-inferiority margin, and data variability (such as standard deviation).

In particular, the sample size calculation in veterinary clinical studies should be based on the primary efficacy parameter, which plays a direct role in assessing the effectiveness of the treatment and provides crucial clinical evidence relevant to the primary objective of the trial. Besides, the specific method for calculating sample size and statistical power may vary depending on the aims of the study, such as assessing superiority (new treatment is better than the current one or placebo), equivalence (two treatments are equivalent within a pre-determined margin), or non-inferiority (new treatment is not worse than the current one within a pre-determined margin).

Additionally, the choice of sample size calculation methods may differ based on the experimental units. In a veterinary clinical study, the experimental unit refers to either an individual animal or the smallest group of animals that can be randomly assigned to a particular treatment. Each treatment represents a distinct set of applied conditions, such as a specific vaccination or other environmental factors (Shere, 2014).

Furthermore, the choice of sample size calculation methods may differ based on the types of variables being investigated to match the statistical methods used in the study. For numerical variables that follow normal distribution (such as number of ticks, average daily weight gain), parametric models (e.g., ANOVA) are often used to compare the mean values; For numerical variables that do not follow a normal distribution (such as clinical score) and time-to-event data (such as survival rates), non-parametric methods (e.g., Wilcoxon rank-sum test) are often used to compare the ranks; For categorical variables (e.g., vomiting vs. non-vomiting), analysis of contingency tables (e.g., Fisher's exact test) is often used to compare the frequency distributions.

Another critical factor in sample size calculation is the expected variability or dispersion of the data within the population being studied. Greater variability usually necessitates a larger sample size to achieve reliable and significant results. The desired power of the study, commonly set at 80 % or 90 % (Martinez-Mesa et al., 2014), indicates the probability of correctly rejecting the null hypothesis when it is false. This, coupled with the expected minimum clinically meaningful effect, helps calculate the sample size needed to detect a difference that is of practical relevance. For a more comprehensive understanding and detailed formulas, beginners are recommended to refer to "Sample Size Calculations in Clinical Research" (Chow et al., 2017; Ko and Lim, 2021) which provides an extensive overview of statistical methods and some practical applications.

Currently, a variety of software tools are available to assist researchers in estimating study power and sample size, including both commercial and free options. Particularly, some of the software are general-purpose statistical tools with sample size determination procedures. The SAS System® (Institute, Cary, North Carolina, USA) is one of the commonly utilized software for calculating sample size, particularly known for its POWER and GLMPOWER procedures. Additionally, R software, with its "pwr," "TrialSize," "G*Power," "powerMediation," "ssize.fdr," and "PowerTOST" packages, provides a free and open-source alternative. IBM SPSS Statistics and STATA Statistical Software (College Station, TX: StataCorp LLC) provide robust tools and functions for sample size determination. Moreover, some Web-based methods like Win Epi (http://www.winepi.net/uk/index.htm) are also gaining adoption. Our review further identifies the use of specialized software such as nQuery, PASS software (NCSS, Kaysville, UT), and Win Episcope 2.0 (CLIVE, Edinburgh, UK). While a plethora of commercial and free software is available, it is crucial for researchers to ensure the accuracy and validity of the sample size calculations. Researchers should verify that the algorithms or formulas provided in the software or websites are accurate and align with the specific objectives of the study.

4.4. Sample size calculation in veterinary clinical trials and in human clinical trials

While the fundamental concepts and statistical methods for sample size calculation, as specified in Dhulkhed et al. (Dhulkhed et al., 2008), are identical for both animal and human clinical trials, the application of these principles differs due to essential variations in experimental designs and regulatory requirements. Animal clinical studies, particularly randomized controlled trials in field studies, often use simpler trial designs compared to human phase 3 clinical trials. This is due to shorter study durations, less complicated endpoints, and fewer ethical constraints. The most common design in animal clinical trials, even at the confirmatory stage, is the parallel group design (Committee for Medicinal Products for Veterinary use, 2021). In contrast, human clinical studies, especially phase 3 trials, frequently require more advanced designs, such as adaptive designs or stepped-wedge cluster designs. Besides, the experimental unit also differs significantly between animal and human trials. In human medicine, the experimental unit is typically an individual patient. In veterinary studies, however, the experimental unit may be a group of animals housed together in a tank, cage, litter, or pen.

Adhering to regulatory guidance is crucial in both human and animal trials to ensure that studies meet ethical standards and generate reliable findings. Organizations such as the FDA, EMA, ICH, and USDA provide guidelines for sample size calculation in human clinical trials, emphasizing statistical adequacy and ethical considerations. Human clinical studies, especially those adhering to rigorous standards like CONSORT (Schulz et al., 2010), require clear justifications for the choice of sample size. Moreover, in some guidelines (EMA, ICH Topic E 9 Statistical Principles for Clinical Trials) the method by which the sample size is calculated should be clearly outlined in the protocol, along with the estimates of any quantities used in these calculations, such as variances, mean values, response rates, event rates, and the differences to be detected. In contrast, animal clinical trials often do not have strict rules for reporting all the details of sample size calculations. This is not surprising as in animal clinical trials, standardized reporting guidelines comparable to those used in human clinical trials are not yet widely published and applied. The two most famous reporting guidelines are REFLECT (an adapted CONSORT-like guideline for livestock) (O'Connor et al., 2010a; O'Connor et al., 2010b) and PetSORT for client- and shelter-owned dog and cat trials (Ruple et al., 2023). However, to extend and promote the usage of REFLECT and PetSORT as CONSORT is difficult because veterinary studies are way less standardized than human studies in many aspects. This standardization is also challenging because of the greater variability in disease types, subjects (e.g., large animal, small animal), and data collection process compared to those in human medicine. Even though, more attention has been paid by veterinary authorities in recent years to require clearer evidence about sample size calculations. To optimize the rigor and reliability of animal trials, it is recommended that future veterinary studies adopt more precise reporting standards. This includes comprehensive justifications for sample size calculation and detailed documentation of statistical methods.

4.5. Strengths and limitations of our review

To the best of our knowledge, the current study includes the most up-to-date literature review of sample size determination in veterinary studies. It offers some theoretical and practical insights, along with recommendations on the appropriate application of sample size determination for various stakeholders in the field of veterinary medicine. This includes, but is not limited to, statisticians, epidemiologists, Research and Development scientists, regulatory affairs managers, and government officials.

One of the limitations of our review is the exclusion of studies that did not use individual animals as the experimental unit. When calculating the sample size for studies with groups of animals as the experimental unit, special attention is needed. This is because the sample size calculated using the same method described in this paper for individual animals then becomes the number of cages for fish or litters for pigs. Additionally, we did not include any study published in languages other than English. This could lead to an overrepresentation of certain viewpoints or findings that are prevalent only in English-language literature.

5. Conclusion

To conclude, our paper not only included the recent evidence about sample size calculation from top animal health journals in terms of a descriptive literature review, but also provided the theoretical and practical viewpoints from experienced statisticians in animal health pharmaceutical industry. Calculating the sample size is one of the first and most crucial steps in planning a veterinary study because the estimated sample size is not an absolute truth, but rather the best approximation. By carefully considering all the critical factors involved in sample size determination and thoroughly documenting the methods, the validity, reliability, and robustness of animal trial findings will be significantly enhanced in future research in animal health.

Funding information

This review was funded by Merck Animal Health, Rahway, New Jersey, USA.

CRediT authorship contribution statement

Zikun Wang: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Qi Cao:** Writing – original draft, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Qingzhi Liu:** Writing – review & editing, Validation, Methodology, Data curation. **Divine Dufe:** Writing – review & editing, Validation, Methodology, Investigation. **Pieter Wouters:** Writing – review & editing, Validation, Methodology, Investigation. **Lijuan Deng:** Writing – review & editing, Validation, Methodology, Investigation. **Annpey Pong:** Writing – review & editing, Validation, Supervision, Methodology, Investigation.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

The work was sponsored by Merck Animal Health, Rahway, New Jersey, USA.

References

Almawly, J., Prattley, D., French, N.P., Lopez-Villalobos, N., Hedgespeth, B., Grinberg, A., 2013. Utility of halofuginone lactate for the prevention of natural cryptosporidiosis of calves, in the presence of co-infection with rotavirus and *Salmonella typhimurium*. Vet. Parasitol. 197 (1–2), 59–67. https://doi.org/10.1016/j.vetpar.2013.04.029.

Arsenakis, I., Michiels, A., Del Pozo Sacristan, R., Boyen, F., Haesebrouck, F., Maes, D., 2017. Mycoplasma hyopneumoniae vaccination at or shortly before weaning under field conditions: a randomised efficacy trial. Vet. Rec. 181 (1), 19. https://doi.org/10.1136/vr.104075.

Arsenakis, I., Boyen, F., Haesebrouck, F., Maes, D.G.D., 2018. Autogenous vaccination reduces antimicrobial usage and mortality rates in a herd facing severe exudative epidermitis outbreaks in weaned pigs. Vet. Rec. 182 (26), 744. https://doi.org/10.1136/vr.104720.

Burden, D.J., Bartley, D.J., Besier, R.B., Claerebout, E., Elliott, T.P., Hoglund, J., Rehbein, S., Torres-Acosta, J.F.J., Van Wyk, J.A., Yazwinski, T., 2024. World Association for the Advancement of Veterinary Parasitology (W.A.A.V.P.): Third edition of the guideline for evaluating efficacy of anthelmintics in ruminants (bovine, ovine, caprine). Vet. Parasitol. 329, 110187. https://doi.org/10.1016/j.vetpar.2024.110187.

Chatzis, M.K., Psemmas, D., Papadopoulos, E., Navarro, C., Saridomichelakis, M.N., 2017. A field trial of a fixed combination of permethrin and fipronil (Effitix((R))) for the treatment and prevention of flea infestation in dogs living with sheep. Parasit. Vectors 10 (1), 212. https://doi.org/10.1186/s13071-017-2145-1.

Chow, S.-C., Shao, J., Wang, H., Lokhnygina, Y., 2017. Sample Size Calculations in Clinical Research, 3rd edn. Chapman and Hall/CRC.

Committee for Medicinal Products for Veterinary use, 2021. Guideline on statistical principles for clinical trials for veterinary medicinal products (pharmaceuticals). European Medicines Agency. Available at: https://assets-us-01.kc-usercontent.com/49d837e7-4c3e-0071-59f0-cb4344fbf121/3755b75b-69a9-488b-b11a-1c444af4802c/guideline-statistical-principles-clinical-trials-veterinary-medicinal-products-pharmaceuticals-rev-1_en.pdf.CVMP.

Dantas-Torres, F., Capelli, G., Giannelli, A., Ramos, R.A.N., Lia, R.P., Cantacessi, C., de Caprariis, D., De Tommasi, A.S., Latrofa, M.S., Lacasella, V., Tarallo, V.D., Di Paola, G., Qurollo, B., Breitschwerdt, E., Stanneck, D., Otranto, D., 2013. Efficacy of an imidacloprid/flumethrin collar against fleas, ticks and tick-borne pathogens in dogs. Parasit. Vectors 6. https://doi.org/10.1186/1756-3305-6-245.

Del Pozo Sacristan, R., Sierens, A., Marchioro, S.B., Vangroenweghe, F., Jourquin, J., Labarque, G., Haesebrouck, F., Maes, D., 2014. Efficacy of early mycoplasma hyopneumoniae vaccination against mixed respiratory disease in older fattening pigs. Vet. Rec. 174 (8), 197. https://doi.org/10.1136/vr.101597.

Dhulkhed, V.D., Mane, Rajesh, Gogate, Vandana, Dhulkhed, Pavan, 2008. Basic Statistical Concepts for Sample Size Estimation. Medknow Publications & Media Pvt. Ltd. (Accessed. European Pharmacopoeia Online (2024): European Directorate for the Quality of Medicines & HealthCare (EDQM). Available at: https://pheur.edqm.eu/home

European Directorate for the Quality of Medicines and HealthCare, 2016. 05 Guideline - EMA Guideline on requirements for the production and control of immunological veterinary medicinal products (Adopted 8 December 2016). European Pharmacopoeia (Ph. Eur.). Available at: https://www.edqm.eu/en/d/95430?p_l_back_url=%2Fen%2Fsearch%3Fq%3Dveterinary%26sort%3Dtitle.

Federal Law Gazette I, p, 1998. Animal Welfare Act (Germany). Available at: https://www.animallaw.info/statute/germany-cruelty-german-animal-welfare-act. Accessed: October 16.

Flecknell, P., 2002. Replacement, reduction and refinement. ALTEX 19 (2), 73–78. https://www.ncbi.nlm.nih.gov/pubmed/12098013.

Forster, S.L., Real, T., Doucette, K.P., King, S.B., 2018. A randomized placebo-controlled trial of the efficacy and safety of a terbinafine, florfenicol and betamethasone topical ear formulation in dogs for the treatment of bacterial and/or fungal otitis externa. BMC Vet. Res. 14 (1), 262. https://doi.org/10.1186/s12917-018-1589-7.

Geurden, T., Smith, E.R., Vercruysse, J., Yazwinski, T., Settje, T., Nielsen, M.K., 2022. World association for the advancement of veterinary parasitology (WAAVP) guideline for the evaluation of the efficacy of anthelmintics in food-producing and companion animals: general guidelines. Vet. Parasitol. 304, 109698. https://doi.org/10.1016/j.vetpar.2022.109698.

Ghasemi, M., Dehpour, A.R., 2009. Ethical considerations in animal studies. J Med Ethics Hist Med 2, 12. Available at: https://www.ncbi.nlm.nih.gov/pubmed/23908726.

Goodman, D., Deis, L., 2005. CiteScore: Web of Science and Scopus. Available at: https://www.scopus.com/sources.uri?RN_AG_Sourced_400000654. Accessed: May 23.

Gupta, K.K., Attri, J.P., Singh, A., Kaur, H., Kaur, G., 2016. Basic concepts for sample size calculation: critical step for any clinical trials! Saudi J Anaesth 10 (3), 328–331. https://doi.org/10.4103/1658-354X.174918.

Hanks, J.E., Larsen, J.W.A., Campbell, A.J.D., 2021. The effect of pasture molluscicide on small lungworm infections and the productivity of grazing lambs. Vet. Parasitol. 292. DOI: ARTN 109395.

Jourquin, S., Bokma, J., De Cremer, L., van Leenen, K., Vereecke, N., Pardon, B., 2022. Randomized field trial comparing the efficacy of florfenicol and oxytetracycline in a natural outbreak of calf pneumonia using lung reaeration as a cure criterion. J. Vet. Intern. Med. 36 (2), 820–828. https://doi.org/10.1111/jvim.16348.

Julious, S.A., 2005. Sample size of 12 per group rule of thumb for a pilot study. Pharm. Stat. 4 (4), 287–291. https://doi.org/10.1002/pst.185.

Kirby, A., Gebski, V., Keech, A.C., 2002. Determining the sample size in a clinical trial. Med. J. Aust. 177 (5), 256–257. https://doi.org/10.5694/j.1326-5377.2002.tb04759.x.

Kneipp, M., Green, A.C., Govendir, M., Laurence, M., Dhand, N.K., 2023. A randomised control trial to evaluate the effectiveness of a commercial vaccine for pinkeye in Australian beef cattle. Prev. Vet. Med. 210. https://doi.org/10.1016/j.prevetmed.2022.105815.

Ko, M.J., Lim, C.Y., 2021. General considerations for sample size estimation in animal study. Korean J. Anesthesiol. 74 (1), 23–29. https://doi.org/10.4097/kja.20662.

Korpivaara, M., Laapas, K., Huhtinen, M., Schoning, B., Overall, K., 2017. Dexmedetomidine oromucosal gel for noise-associated acute anxiety and fear in dogs-a randomised, double-blind, placebo-controlled clinical study. Vet. Rec. 180 (14), 356. https://doi.org/10.1136/vr.104045.

Korpivaara, M., Huhtinen, M., Aspegrén, J., Overall, K., 2021. Dexmedetomidine oromucosal gel reduces fear and anxiety in dogs during veterinary visits: a randomised, double-blind, placebo-controlled clinical pilot study. Vet. Rec. 189 (12). https://doi.org/10.1002/vetr.832.

Larsen, I., Baekbo, P., Nielsen, J.P., 2023. Umbilical Outpouchings in two Danish herds raising pigs with minimal use of antibiotics-Results from a field trial evaluating the efficacy of autogenous vaccines and iodine application. Prev. Vet. Med. 214. DOI: ARTN 105905.

Lascelles, B.D., Rausch-Derra, L.C., Wofford, J.A., Huebner, M., 2016. Pilot, randomized, placebo-controlled clinical field study to evaluate the effectiveness of bupivacaine liposome injectable suspension for the provision of post-surgical analgesia in dogs undergoing stifle surgery. BMC Vet. Res. 12 (1), 168. https://doi.org/10.1186/s12917-016-0798-1.

Martinez-Mesa, J., Gonzalez-Chica, D.A., Bastos, J.L., Bonamigo, R.R., Duquia, R.P., 2014. Sample size: how many participants do I need in my research? An. Bras. Dermatol. 89 (4), 609–615. https://doi.org/10.1590/abd1806-4841.20143705.

Masmeijer, C., Devriendt, B., Rogge, T., van Leenen, K., De Cremer, L., Van Ranst, B., Deprez, P., Cox, E., Pardon, B., 2019. Randomized field trial on the effects of body weight and short transport on stress and immune variables in 2- to 4-week-old dairy calves. J. Vet. Intern Med. 33 (3), 1514–1529. https://doi.org/10.1111/jvim.15482.

Masset, N., Meurens, F., Marie, M., Lesage, P., Lehébel, A., Brisseau, N., Assié, S., 2020. Effectiveness of two intranasal vaccines for the control of bovine respiratory disease in newborn beef calves: a randomized non-inferiority multicentre field trial. Vet. J. 263. https://doi.org/10.1016/j.tvjl.2020.105532.

Meadows, C., Guerino, F., Sun, F., 2014. A randomized, blinded, controlled USA field study to assess the use of fluralaner tablets in controlling canine flea infestations. Parasit Vectors 7.

Nixon, S.L., Rose, L., Muller, A.T., 2019. Efficacy of an orally administered anti-diarrheal probiotic paste (Pro-Kolin Advanced) in dogs with acute diarrhea: a randomized, placebo-controlled, double-blinded clinical study. J. Vet. Intern. Med. 33 (3), 1286–1294. https://doi.org/10.1111/jvim.15481.

O'Connor, A.M., Sargeant, J.M., Gardner, I.A., Dickson, J.S., Torrence, M.E., Consensus Meeting, P., Dewey, C.E., Dohoo, I.R., Evans, R.B., Gray, J.T., Greiner, M., Keefe, G., Lefebvre, S.L., Morley, P.S., Ramirez, A., Sischo, W., Smith, D.R., Snedeker, K., Sofos, J., Ward, M.P., Wills, R., 2010a. The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety by modifying the CONSORT statement. Zoonoses Public Health 57 (2), 95–104. https://doi.org/10.1111/j.1863-2378.2009.01311.x.

O'Connor, A.M., Sargeant, J.M., Gardner, I.A., Dickson, J.S., Torrence, M.E., Dewey, C. E., Dohoo, I.R., Evans, R.B., Gray, J.T., Greiner, M., Keefe, G., Lefebvre, S.L., Morley, P.S., Ramirez, A., Sischo, W., Smith, D.R., Snedeker, K., Sofos, J., Ward, M. P., Wills, R., Steering, C., 2010b. The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety. J. Vet. Intern. Med. 24 (1), 57–64. https://doi.org/10.1111/j.1939-1676.2009.0441.x.

Otranto, D., Dantas-Torres, F., Fourie, J.J., Lorusso, V., Varloud, M., Gradoni, L., Drake, J., Geurden, T., Kaminsky, R., Heckeroth, A.R., Schunack, B., Pollmeier, M., Beugnet, F., Holdsworth, P., 2021. World Association for the Advancement of veterinary parasitology (W.A.A.V.P.) guidelines for studies evaluating the efficacy of parasiticides in reducing the risk of vector-borne pathogen transmission in dogs and cats. Vet. Parasitol. 290, 109369. https://doi.org/10.1016/j.vetpar.2021.109369.

Paré, G., Jaana, M., Kitsiou, S., 2015. Synthesizing information systems knowledge: a typology of literature reviews. Inf. Manag. 52 (2), 183–199. https://doi.org/10.1016/j.im.2014.08.008.

Pithua, P., Aly, S.S., Haines, D.M., Champagne, J.D., Middleton, J.R., Poock, S.E., 2013. Efficacy of feeding a lacteal-derived colostrum replacer or pooled maternal colostrum with a low IgG concentration for prevention of failure of passive transfer in dairy calves. Javma-Journal of the American Veterinary Medical Association 243 (2), 277–282. https://doi.org/10.2460/javma.243.2.277.

Ruple, A., Sargeant, J.M., Selmic, L.E., O'Connor, A.M., 2023. The standards of reporting randomized trials in pets (PetSORT): methods and development processes. Front Vet Sci 10, 1137774. https://doi.org/10.3389/fvets.2023.1137774.

Salichs, M., Badiella, L., Sarasola, P., Homedes, J., 2022. Efficacy and safety of enflicoxib for treatment of canine osteoarthritis: a 6-week randomised, controlled, blind, multicentre clinical trial. Vet. Rec. 191 (6), e949. https://doi.org/10.1002/vetr.949.

Sandelin, A., Härtel, H., Seppä-Lassila, L., Kaartinen, L., Rautala, H., Soveri, T., Simojoki, H., 2020. Field trial to evaluate the effect of an intranasal respiratory vaccine protocol on bovine respiratory disease incidence and growth in a

commercial calf rearing unit. BMC Vet. Res. 16 (1). https://doi.org/10.1186/s12917-020-02294-7.

Saridomichelakis, M.N., Chatzis, M.K., Petanides, T., Papadopoulos, E., 2015. A field trial of spinosad for the treatment and prevention of flea infestation in shepherd dogs living in close proximity to flea-infested sheep. Parasit Vectors 8, 324. https://doi.org/10.1186/s13071-015-0945-8.

Schoster, A., Staempfli, H.R., Abrahams, M., Jalali, M., Weese, J.S., Guardabassi, L., 2015. Effect of a probiotic on prevention of diarrhea and Clostridium difficile and Clostridium perfringens shedding in foals. J. Vet. Intern. Med. 29 (3), 925–931. https://doi.org/10.1111/jvim.12584.

Schulz, K.F., Altman, D.G., Moher, D., Grp, C., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Bmj-British Medical Journal 340. https://doi.org/10.1136/bmj.c332.

Serdar, C.C., Cihan, M., Yucel, D., Serdar, M.A., 2021. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochem Med (Zagreb) 31 (1), 010502. https://doi.org/10.11613/BM.2021.010502.

Shere, J., 2014. U.S. Department of Agriculture Animal and Plant Health Inspection Service Veterinary Services (2014) *General Licensing Considerations: Efficacy Studies for Prophylactic. and Therapeutic Biologics*. Washington, D.C.

Singh, N., Vayer, P., Tanwar, S., Poyet, J.-L., Tsaioun, K., Villoutreix, B.O., 2023. Drug discovery and development: introduction to the general public and patient groups. Frontiers in Drug Discovery 3. https://doi.org/10.3389/fddsv.2023.1201419.

Suokko, M., Saloranta, L., Lamminen, T., Laine, T., Elliott, J., 2020. Ropinirole eye drops induce vomiting effectively in dogs: a randomised, double-blind, placebo-controlled clinical study. Veterinary Record 186 (9). https://doi.org/10.1136/vetrec-2018-104953.

Wareham, K.J., Hyde, R.M., Grindlay, D., Brennan, M.L., Dean, R.S., 2017. Sample size and number of outcome measures of veterinary randomised controlled trials of pharmaceutical interventions funded by different sources, a cross-sectional study. BMC Vet. Res. 13 (1), 295. https://doi.org/10.1186/s12917-017-1207-0.