Review

THE VALUE OF N-OF-1 DATA IN ZOOLOGICAL MEDICINE: A METHODOLOGICAL REVIEW

Charles O. Cummings[1,*], Jessica Eisenbarth [2], and David D.R. Krucik [3]

1 Tufts Clinical and Translational Science Institute, Tufts Medical Center, Boston, MA, USA 02111

2 Department of Clinical Sciences, Cummings School of Veterinary Medicine at Tufts University, North Grafton, MA, USA 01536

3 Department of Comparative Medicine, Stanford University, Stanford, CA, USA 94305

* Correspondence: cummings3793@gmail.com

Abstract : N-of-1 data are unavoidable in zoological medicine. Accordingly, zoological medicine clinicians and investigators need research techniques that can make use of these data. This article reviews two methodologies for using both observational and experimental N-of-1 data: 1) systematic reviews and meta-analyses of case reports and 2) prospective N-of-1 trials. Systematic reviews of case reports and other observational evidence are formal, unbiased summaries the clinical characteristics of a particular disease-taxon combination. They offer advantages to narrative reviews by minimizing omission of relevant articles, thereby reducing the potential for mischaracterization of the literature. Meta-analyses are extensions of systematic reviews that quantitatively synthesize the data from the included articles. While valuable, systematic reviews and meta-analyses of case reports can have limited interpretations due to publication bias and confounding present in their source materials. In contrast to case reports, N-of-1 trials are prospective study designs that allow clinicians to make strong inferences about the effect of an intervention in a particular patient. They are double-blinded, single patient, multi-crossover studies that are of particular value in fields where it is difficult to recruit sufficient patients for

 doi:10.20944/preprints202212.0424.v1

conventional randomized control trials (RCTs), such as zoological medicine. Because they require multiple crossover periods, N-of-1 trials are ideal for evaluating short-acting interventions in patients with somewhat stable chronic diseases, such as osteoarthritis. More complex than conventional therapeutic trials, N-of-1 trials require prior consideration of how to achieve blinding, appropriate placebo controls, quantitative primary outcomes, analysis methods, and ethical approval. Aggregation of N-of-1 trials allows estimation of the average treatment effect across the population with fewer participants than a conventional RCT. While systematic reviews and meta-analyses of case reports can be used to synthesize the observational N-of-1 data already in existence, N-of-1 trials offer an exciting way to prospectively generate strong evidence that will be useful for evidence-based decision-making.

Keywords: systematic review; meta-analysis; evidence synthesis; RCT; N-of-1 trial; single subject design

---

1. Introduction

Evidence-based medicine (EBM) is the "conscientious, explicit, judicious, and reasonable use of modern, best evidence in making decisions about the care of individual patients."[19] This best evidence is typically determined according to the study design of individual studies. Generally speaking, randomized controlled trials (RCTs) – including RCTs within a single patient, or N-of-1 trials – are considered higher quality evidence than cohort studies which are superior to case-control studies, case series, and N-of-1 case reports.[25] Importantly, methodological quality is also weighed when making determinations about evidence quality; for example, a well conducted cohort study that demonstrates a large effect size could be considered superior to a poorly designed, underpowered RCT. Systematic reviews and meta-analyses are sometimes placed at the apex of evidence quality pyramids, but some authors prefer to see them as a lens with which evidence is viewed, because their value is dependent on the studies they include.[25] A systematic review and meta-analysis of well-conducted RCTs results in superior evidence to a systematic review and meta-analysis of case reports and case series. Likewise, one good quality RCT could

 doi:10.20944/preprints202212.0424.v1

be better evidence of an intervention than a meta-analysis of several case reports/series.

In zoological medicine, the resources for RCTs (cases, money, etc.) are dwarfed by the myriad of diseases and species that a clinician may have to treat. As a result, clinicians might only have descriptive data from case reports/series or RCTs from distantly related taxa to aid in their decision-making. In this article, we will review two methodologies for using the N-of-1 data that are frequently encountered in zoological medicine: 1) systematic review and meta-analysis of case reports and 2) prospective N-of-1 clinical trials.

2. Observational N-of-1 Data Background

Similar to human medicine, case reports in zoological medicine are increasingly passed over in favor of publishing experimental trials and larger-scale observational studies.[15,27] This inclination towards publishing higher tiers of evidence is understandable, but the utility and worthiness of case reports – including multiple case reports on the same disease-taxon combination – should not be discounted. Case reports can describe new phenotypes or genotypes of a disease, emerging pathogens, common manifestations of rare diseases, rare manifestations of common diseases, adverse drug effects, and novel treatments.[26] They can also be illustrative for teaching others to avoid pitfalls in case management.[6,26] Given this diversity of roles, it is not surprising that some case reports actually become highly influential pieces. In fact, when the American Medical Association reprinted 51 seminal articles in its journal in 1985, case reports made up 10% of those selected.[11,26] Additionally, many of the *Journal of Zoo and Wildlife Medicine*'s (JZWM) most purchased articles are case reports (Lung, pers. comm.).

In zoological medicine, there is often a race to first publication on a disease-taxon combination, but subsequent N-of-1 case reports that further describe the condition may be lacking because they are viewed as no longer novel. In addition, some editors have become hesitant to publish any N-of-1 case reports (Lung, pers. comm.), possibly because they contain preliminary observations that may later be found to be untrue or misleading.[3] Thus, we typically rely on clinicians, institutions, or veterinary advisors to have knowledge of multiple cases of a disease-taxon combination before more information on that disease-taxon combination is published. While a reluctance to publish single case reports is understandable in some fields – a single case report of myocardial infarction in humans is exceedingly unlikely to affect disease understanding

 doi:10.20944/preprints202212.0424.v1

or treatment – there is a growing understanding of the benefit of multiple case reports of rare diseases.[26,36,37] In zoological medicine, clinicians treat hundreds of different conditions in hundreds of different species; thus, it is reasonable to view many of these disease-taxon combinations as rare diseases.

3. Narrative Synthesis of N-of-1 Case Reports

Systematic reviews are one way in which N-of-1 case reports can be used in synthesizing evidence from rare diseases.[26,27,34,36,46] Systematic reviews are a formal, reproducible, and unbiased approach to characterize the literature that pertain to specific questions – Population, Intervention, Comparator, Outcome (PICO) questions – and use specific study designs.[1,30] They report search terms, when searches were performed, the databases and indices searched, as well as inclusion and exclusion criteria.[1,30] The quality of included studies is also evaluated. More detailed information on how to conduct systematic reviews can be found elsewhere.[1,30] Scoping reviews are similar in their systematic approach, but they tend to cover a topic more broadly (as opposed to specific PICO questions) and do not appraise study quality.

Systematic reviews of N-of-1 case reports and larger case series summarize the commonalities and differences between cases to give the reader a more complete clinical picture. In human medicine, these reviews have been used to describe the patient demographic characteristics, presenting signs and symptoms, comorbidities, diagnostic findings, therapeutic management, and outcomes in patients with COVID-19-associated myocarditis and vasculitis, mucopolysaccharidosis, and more.[34,36,46] In zoological medicine, scoping reviews and systematic reviews have been used to describe a wide breadth of topics.[5,33,42] In an article describing clinical presentation and management of snakebites in birds, the avian and snake species involved, bite locations, clinical signs, treatments, outcomes, and adverse events were all summarized.[5] Approximately 10% of the 31 records included in that review were N-of-1 case reports.[5] Other authors have used case reports in combination with other morbidity and mortality reviews to document the causes of morbidity and mortality amongst captive great apes.[42] In yet another study, this time of surgical outcomes among captive and free-ranging wild animals, the authors systematically reviewed 635 abstracts, of which 60% were single case reports, and were able to summarize surgical complication rate by taxa: 4.7% for mammals, 3.7% for birds, and 7.7% for

 doi:10.20944/preprints202212.0424.v1

reptiles. Notably, however, when these results were compared to a comprehensive review of surgical procedures at major zoological institutions in the United Kingdom, they were found to markedly underestimate the surgical complication rate for both birds and for reptiles (mammals were not evaluated).[33] This discrepancy may be a result of several factors but the most notable one is publication bias among case reports. Publication bias is when the direction or magnitude of a study's results affects the decision to publish.[29] For case reports, this suggests an inclination to preferentially publish cases that had positive outcomes.[29]

4. Quantitative Synthesis of N-of-1 Case Reports

A meta-analysis adds to a systematic review by providing a quantitative synthesis of information from that review. Not all systematic reviews will have a meta-analysis, but all meta-analyses should be associated with a systematic review. The goal of these analyses is to provide an estimate for a treatment effect with greater power and precision than what may be presented in any one of the individual studies. Traditional meta-analyses use data from RCTs to estimate a pooled treatment effect (e.g., odds ratio, risk ratio, mean weighted difference) for a specific intervention. To do this, data from as few as two studies may be used to create a forest plot, which shows graphically the treatment effect estimate and confidence intervals from each included study as well as a new pooled estimate. While non-randomized studies cannot be used to estimate an unbiased treatment effect, they can be used to estimate proportions, e.g., the prevalence of a certain clinical sign in all individuals with a particular disease.[2]

One novel variation on the proportional meta-analysis that has been used in zoological medicine is to pool patients of the same taxa from different case reports or case series. In one such meta-analysis, investigators aimed to understand differences in complication rates in wildlife undergoing open surgery and minimally invasive surgery (MIS). Unfortunately, there were only two small studies that directly compared complication rates of open vs. MIS in the same species for the same indication; meta-analysis of these studies favored MIS, but not significantly. The authors then aggregated cases from the literature and summarized the complication rate of MIS vs. open surgery in birds, mammals, reptiles, and fish using risk ratios (i.e., complication rate in MIS/complication rate in open surgery; Figure 1). These taxa-specific risk ratios were then used

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0424.v1

to estimate an overall pooled risk ratio of 0.21 (95% CI: 0.15-0.30; Figure 1), meaning that across all taxa the MIS complication rate was roughly one fifth that of open surgery.

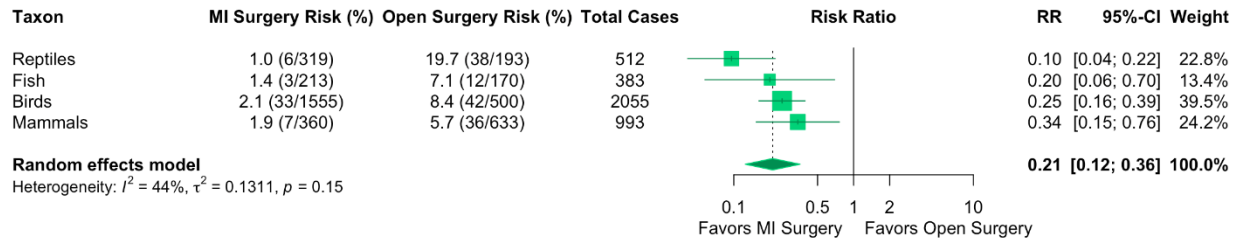

Figure 1. Example of a forest plot, generated using data from a previous systematic review and meta-analysis.[33] Data on surgical complication rates in captive and free-ranging wildlife were aggregated to the taxon levels above, complication rates of both minimally invasive (MI) surgery and open surgery were calculated, followed by calculation of risk ratios (squares). A pooled risk ratio across all taxa was also calculated (diamond) using a random effects model. Importantly, the data used to generate this plot were sourced largely from non-comparative studies which means that confounding significantly limits interpretation of the plot; this plot is simply to demonstrate how forest plots appear.

Importantly, the authors noted several limitations to this analysis. The biggest limitation was that the vast majority of included studies were non-comparative and not randomized, which can lead to *considerable* confounding. For example, instead of complication rates being different due to different surgical approaches, complication rates could have been higher in open surgery studies because they were performed more often on less hardy species that tolerate anesthesia poorly. Likewise, animals in the included MIS articles may have been healthy individuals undergoing elective surgery (e.g., sterilization) vs. sick animals undergoing open surgery for a tumor resection or a septic peritonitis. Even taking these significant limitations into consideration, it is reasonable to interpret the above meta-analysis as describing MIS to be an approach that *possibly* results in lower complication rates on a macrolevel, but that recommendations for minimally invasive over open approaches for any particular clinical scenarios cannot be made without more specific evidence. These types of conclusions, while vague, may be the best that can be achieved until more evidence from experimental studies and trials accumulates.

Lastly, individual patient data can be compiled from N-of-1 case reports and case series into a single dataset to perform other types of statistical analysis beyond the meta-analyses described

 doi:10.20944/preprints202212.0424.v1

above (e.g., hypothesis testing and/or estimation using frequentist or Bayesian approaches).[26] In human medicine, linear regression analyses have been performed using case reports of cardiac dysfunction and hypocalcemia to show a significant link between left ventricular ejection fraction and corrected total serum calcium concentration.[26,28] In another quantitative synthesis, the authors used t-tests to assess differences in mortality for 143 aortic intramural hematoma patients that were treated surgically vs. medically.[17] In zoological medicine, similar analyses are starting to appear in the literature. One such analysis reviewed cases from the amphibian neoplasia literature to supplement the Exotic Species Cancer Research Alliance (ESCRA) database.[13] In that analysis, the authors used boosted regression modeling to identify patient, tumor, and management factors associated with positive outcomes in 50 cases.[13] Notably, these cases included 28 different species and 15 different tumor types.[13] While these types of analyses in zoological medicine are promising, it may be best to view the conclusions from these quantitative syntheses as tentative or exploratory/hypothesis-generating more so than establishing definitive links. This is because often data from very different species, conditions, and treatments might have been combined. Possible exceptions to this notion include disease-taxon combinations that are particularly well described in the literature with large numbers of case reports and case series, like elephant endotheliotropic herpesviruses (EEHV) in Asian elephants (*Elephas maximus*).

5. Limitations of Synthesizing Observational N-of-1 Data

As discussed previously, systematic reviews and meta-analyses are tools with which one can assess contemporary evidence on a topic of clinical interest. They do not transcend the biases present in their source material. Many case reports are subject to publication and treatment indication bias, with positive cases possibly more likely to be published than those with worse outcomes. This has the potential to leave useful insights unpublished. Undoubtedly, though, the biggest limitation to synthesizing case reports is that they are not randomized or comparative, which can lead to extensive confounding and limited interpretations.

6. Considerations for Publishing and Synthesizing Observational N-of-1 Data

In order to make clinical observations most useful when published, authors should strongly consider adhering to case report (CARE) guidelines published for human medicine[10]. These guidelines were specifically developed to increase the accuracy, transparency, and usefulness of

 doi:10.20944/preprints202212.0424.v1

case reports.[10] Even in journals that recommend the use of CARE guidelines, reporting is often imperfect and/or incomplete, so it is on would-be evidence synthesizers to appraise case reports for adherence to the CARE guideline checklist.[3] Additionally, authors can consider making raw clinical data (e.g., CBCs and serum/plasma chemistries) available as supporting information to allow for use in quantitative syntheses.

When setting out to perform a systematic review of case reports, it is necessary to consider the databases you intend to search. A single database is typically considered insufficient, but searching multiple databases should provide sufficient coverage to avoid omitting relevant publications and mischaracterizing the literature. MEDLINE (accessed through PubMed), for example, is an excellent database of medical research that should be searched when performing a systematic review. However, if a zoological medicine clinician were to *only* search MEDLINE, they would not find any results from the *Journal of Herpetological Medicine and Surgery* or the *Journal of Exotic Pet Medicine*, among others, as neither are currently indexed by MEDLINE. Useful databases for zoological medicine include MEDLINE, CAB Abstracts, Zoological Record, BIOSIS Citation Index, SciELO Citation Index, among others, as well as Google Scholar, which is technically not a database but useful nonetheless.[5,13,42] Beyond simply finding publications, investigators synthesizing observational N-of-1 data may also want to consider using a proposed tool to appraise methodological quality of case reports and case series.[26] This tool asks questions pertaining to selection (i.e., does this case represent the entirety of the author's experience with the disease-taxon combination?), ascertainment of exposures and outcome, causality, and reporting quality.[26]

7. Observational N-of-1 Data Conclusions

While the use of registries like ESCRA could and should expand in zoological medicine, observational N-of-1 data published as case reports will continue to be valuable contributions to the literature. These case reports are useful for a number of reasons – such as identifying a new disease or showing others a potential pitfall in clinical management – and can be published shortly after the experience instead of waiting a number of years to accrue a critical mass of cases. For insufficiently described disease-taxon combinations, it is reasonable to believe that addition of a single case may impact the average clinician's understanding of the combination, so delay in

 doi:10.20944/preprints202212.0424.v1

publication may be unwise. This being said, the real power of observational N-of-1 data lies in the different ways we can aggregate them to better understand different diseases, diagnostics, and treatment approaches.

8. Experimental N-of-1 Data Background

In contrast to case reports – which often result in findings akin to "we performed this intervention, and the patient survived" – rigidly conducted N-of-1 trials are prospective study designs that allow clinicians to clearly demonstrate the effect of an intervention in an individual patient.[16,44] In fact, N-of-1 trials are really just special cases of RCTs. They feature pre-specified, quantitative outcomes, placebo controls, treatment crossovers, and blinding of both the outcome assessor and the investigator. They even have their own reporting guidelines, CENT (Consort Extension for reporting N-of-1 Trials).[40]

Because N-of-1 trials require crossover from the experimental to the control condition and vice versa, they cannot be used for all diseases. Some conditions, such as infections, will be cured with a short course of treatment while other conditions will result in rapid clinical declines before a crossover can occur. Thus, the conditions best suited for N-of-1 trials are those with stable, chronic courses, such as osteoarthritis, epilepsy, and others. Likewise, N-of-1 trials cannot be used for all interventions. Some interventions will require long periods of treatment to demonstrate an effect (e.g., acupuncture or amantadine), while others may have prolonged effects that extend well beyond the treatment period. Depending on the patient and disease, these types of interventions may require several treatment and/or washout periods that are longer than desired. Thus, the interventions best suited for N-of-1 trials are typically fast-acting and short-lived.

In human medicine, N-of-1 trials are ideal for rare conditions or care settings where it is difficult or impossible to recruit sufficient patients for an RCT.[7,23] Additionally, N-of-1 trials have received increased attention as a result of precision medicine initiatives.[18,38] This is especially true for patients with multiple comorbidities who would not typically be enrolled in RCTs.[23] N-of-1 trials have been used for evaluation of common interventions for chronic pain, migraines, chemotherapy-induced nausea, and depression.[16] They have also been used to assess novel treatments for COPD, asthma, cystic fibrosis, chronic fatigue, and others.[16]

 doi:10.20944/preprints202212.0424.v1

9. Considerations for Performing N-of-1 Trials in Zoological Medicine

There are several factors to consider when planning an N-of-1 trial. These include quantifiable outcomes, run-in periods, blinding, placebo, number and length of intervention periods, length of washout periods, analysis methods, and ethical approval.

Perhaps the most important consideration in an N-of-1 trial is deciding upon the primary study outcome. Ideally, outcomes should be easily quantifiable. These outcomes could derive from laboratory data (e.g., hepatic enzyme concentrations for chronic hepatitis), wearable health monitors (e.g., steps per day for osteoarthritis), or keeper documentation (e.g., journals detailing seizure events for epilepsy). Some of these outcomes may be more precise than others, such as laboratory data vs. keeper-recorded observations. However, for some conditions like epilepsy, keeper-recorded observations may be the best measurement we can achieve, provided these outcomes are recorded consistently and systematically throughout the trial period. Before the trial commences, it is also important to consider what constitutes a clinically meaningful change in the outcome measure.

In RCTs, a run-in period is the time prior to randomization when all participants are on the same intervention.[8] Run-in periods are useful for determining potential compliance difficulties before the trial officially begins; this way, if a participant is unable to, for example, take medications three times daily, they can be excluded prior to being randomized.[8] In the same way, it may be useful to use a run-in period to ensure the intervention is even feasible before planning an entire N-of-1 trial. This run-in period does not have to be particularly formal either. A clinician may simply prescribe a medication/supplement as they would regularly. If the intervention results in a rapid, dramatic improvement in clinical signs, there may be no need for a trial. However, if the intervention is feasible and there is uncertainty about the degree of clinical improvement, an N-of-1 trial might be indicated.

Blinding refers to whether the study participant, the person who assesses the outcome, or the investigator know whether a particular patient is receiving the intervention or the placebo. When trials are double-blinded, it means that neither the patient, the outcome assessor, nor the investigator know.[39] Blinding is important to trials because it reduces the potential for bias.[22] If investigators know which intervention is the true treatment vs. placebo, they may be inclined to

 doi:10.20944/preprints202212.0424.v1

analyze the data in a way that may be more likely to show a favorable result. Likewise, if the outcome assessor, such as a zookeeper, knows that the animal was receiving the treatment and not the placebo this may subconsciously influence their assessments;[24,39] this is akin to a placebo by proxy effect.[45] In human medicine, blinding can be facilitated by the use of clinical trial offices or research pharmacists, but in zoological medicine we typically lack these resources.[39] Instead, investigators may opt to simply have a clinician or technician who is not otherwise involved in outcome assessment or statistical analysis label both placebo and treatment regimens with different random numbers or letters, making sure to record which is treatment and which is placebo to allow unblinding after analysis. Still, there may be some scenarios in which blinding is infeasible or impractical. In such cases, single patient, open trials (SPOTs) may be a beneficial middle ground between formal N-of-1 trials and the informal therapeutic trials commonly used in practice.[41]

The selection of placebo, or sham treatment in the case of non-pharmaceutical interventions, is also important. An ideal placebo should be almost identical in smell, taste, and appearance to the treatment, lacking only the active ingredient.[39] This may require use of a compounding pharmacy to make both the treatment and a placebo that is similar in these aspects.[39] While a perfect placebo may not be achievable, a reasonable attempt should be made so as to prevent accidental unblinding of treatment administrators (often the same people as outcome assessors), especially when commonly used drugs may have recognizable characteristics. In addition to preventing accidental unblinding, appropriate placebo controls should be used because placebo effects have been documented in animals even in double blinded trials;[24] in one meta-analysis of double-blinded, placebo-controlled, canine epilepsy trials, 79% of dogs (22/28) that received placebo had reductions in seizure frequency from baseline.[24]

A crossover period is the time during which the patient may be on either placebo or intervention. N-of-1 trials include multiple crossover periods. The number and length of crossover periods should be based on the nature of the intervention and the study outcomes.[16] Increasing the number and length of crossover periods (as well as the number of measurements within each crossover period) results in greater statistical power to detect to a difference between active treatments and placebo, but they come at the expense of cost and possibly compliance.[32] If

 doi:10.20944/preprints202212.0424.v1

a large treatment effect is expected, then as few as four crossover periods alternating between pairs of treatment and placebo (e.g., ABBA) may be sufficient.[4,32] However, if multiple different treatments or dosing schemes are trialed, the number of crossover periods and length of the trial may rapidly increase. Beyond the number and length of crossover periods, the sequence of crossover periods also warrants consideration. Some authors propose blocked randomization. In blocked randomization, the sequence is randomized within each pair/trio/etc., making sequences such as ABABAB or ABBAB acceptable, but sequences like AAABBB unacceptable. Still, other authors suggest alternating pairs of AB and BA (e.g., ABBA, ABBAAB, etc.) to be optimal or near optimal in most situations with just one active treatment and one placebo.[4,39] In short, N-of-1 trials can be as simple or complex as the investigators wish.

Investigators will also want to consider the possibility of carryover effects.[16,32] This is when the effects of one crossover period continue into another period, potentially altering the results. When this occurs, for example the beneficial effect of a treatment extending into the placebo period, there is a reduction in contrast between the two periods, resulting in a loss of statistical power.[32] To account for this, many crossover trials include washout periods of varying duration, during which neither treatment nor placebo are administered. The exact length of washout periods will be dependent on the intervention. For example, one commonly used heuristic in small animal medicine dictates that clinicians should allow approximately 5-7 days of washout before switching a patient from one NSAID to another. Even with washout periods, however, the possibility of carryover effects cannot be completely excluded so they must be considered during statistical analyses.

While sometimes conclusions on treatment effect can be drawn from visual inspection of the plotted data, investigators often want to be able to quantify treatment effect.[20] To do this, statistical analysis methods must be carefully planned in advance. Given the potential complexity of the data, consultation with a biostatistician familiar with single subject research designs may be useful. Because the data are collected longitudinally in time, investigators must consider autocorrelation, also known as serial correlation, of their data.[16,21,35] Autocorrelation refers to the tendency of data that were collected sequentially in time to be more alike. This is intuitive to clinicians in their clinical practice as they, for example, serially monitor blood chemistry analytes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0424.v1

over the course of several days. If autocorrelation is severe enough, neglecting to take it into account can result in erroneous conclusions, such as a failure to detect a genuine treatment effect.[16,20] More detailed discussion of the analytical considerations needed for N-of-1 trials can be found elsewhere.[9,20]

One final consideration to make is one of ethical approval. N-of-1 trials are at the border between the "experimental treatments" we are faced with every day as zoo clinicians and true "medical scientific research."[43] In human medicine, one proposed framework for deciding whether or not N-of-1 trials require prior formal ethical approval has three considerations. These considerations are summarized as a flow chart in Figure 2. Importantly, these are only intended as guidelines, and we encourage clinicians to consult with their own IACUC or research committees.

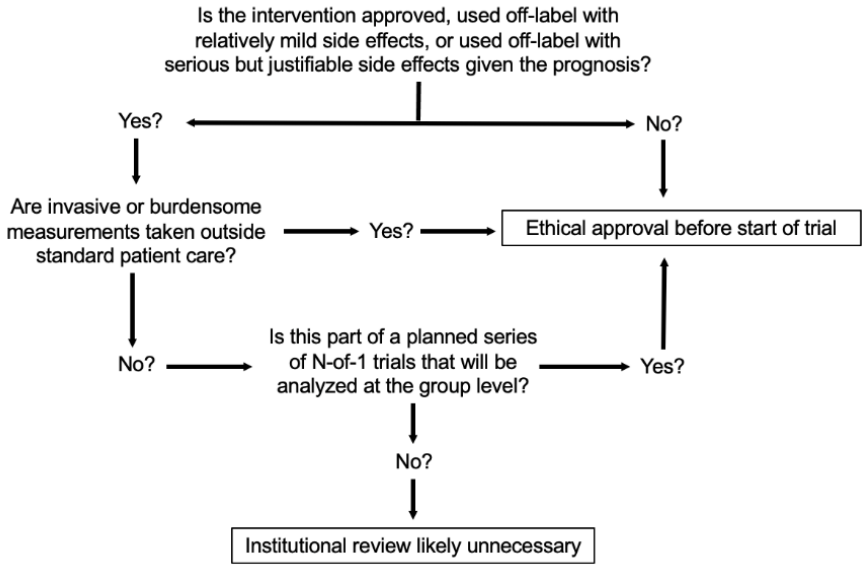

Figure 2. Flowchart, adapted from human medicine,[43] for determining whether a proposed N-of-1 trial is "experimental treatment" or "medical scientific research," which requires prior ethical approval.

10. Potential Applications of N-of-1 Trials in Zoological Medicine

Within zoological medicine, there are almost never species-specific RCTs that can guide therapy because it would be extremely difficult to recruit sufficient patients. Yet, because we must still treat our patients, we sometimes resort to using interventions of unclear mechanistic

 doi:10.20944/preprints202212.0424.v1

appropriateness for a particular taxon (e.g., loop diuretics in reptiles).[31] In such situations and others where there is uncertainty of treatment effect, N-of-1 trials offer clinicians a way to better inform their treatment decisions as well as avoid polypharmacy by discontinuing ineffective therapies.[14] As in human medicine, these trials could be used across a range of chronic diseases, including, but not limited to, osteoarthritis, epilepsy, cardiac disease, hypertension, diabetes, inflammatory/autoimmune diseases, and allergic disease. Additionally, N-of-1 trials may be useful in understanding which drugs or environmental modifications can reduce a particular patient's display of abnormal behaviors and stereotypies, such as feather-picking.

11. Hypothetical N-of-1 Trial Using Simulated Data

To illustrate the process of an N-of-1 trial, consider the following hypothetical study. A clinician was interested in determining whether an oral NSAID would increase the mobility and, in turn, quality of life of an osteoarthritis-affected thylacine (*Thylacinus cynocephalus*). The clinician first decided on the outcome; in this case, they decided daily step count from a collar-mounted, dog movement tracker would be a good, objective, primary outcome. Because NSAIDs are not necessarily benign, the clinician then set out to determine what was the minimum treatment effect that warranted NSAID continuation after the trial. After consultation with keeper staff and other stakeholders, the clinician decided that a 25% increase in mean daily step count would indicate that the benefits outweighed the potential complications of treatment. This way, they would avoid "moving the goalposts" if the NSAID only resulted in some smaller mobility benefit, even if that benefit was statistically significant.

In discussions of study design, keeper staff preferred a simple, four crossover period, ABBA design (Figure 3) with each crossover period consisting of 7 days on either treatment or placebo followed by a 2 day washout period. This was preferred because it would be shorter and thus would allow the thylacine to receive consistent treatment sooner if it was effective. The clinician, however, was not sure and preferred more crossover periods so that they would be more likely to get the "correct" answer. They eventually agreed to a six crossover period, ABBAAB design (Figure 3). With the design settled, the clinician reached out to a compounding pharmacy to make up both NSAID and placebo in identical formulations. A veterinary technician uninvolved in the animal's care then relabeled the drug bottles, "A" and "B", and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2022 doi:10.20944/preprints202212.0424.v1

recorded their identity in a sealed envelope. The trial was then carried out according to plan, and, at the end of the trial, the step count data was downloaded from the collar.

When the clinician plotted the data (Figure 3), it was apparent from visual inspection alone that the thylacine's mobility improved while on treatment A. However, the primary question was "did the step counts increase by 25% when on an NSAID compared to placebo?" Thus, the clinician calculated the mean daily step counts of treatment A (20,599 steps) and treatment B (13,958), an almost 50% increase in steps. The clinician did not feel the need to perform any further statistical analysis, as they realized that if treatment A was the NSAID, they would continue it whether statistical significance was achieved or not. After unblinding, treatment A was, in fact, the NSAID.

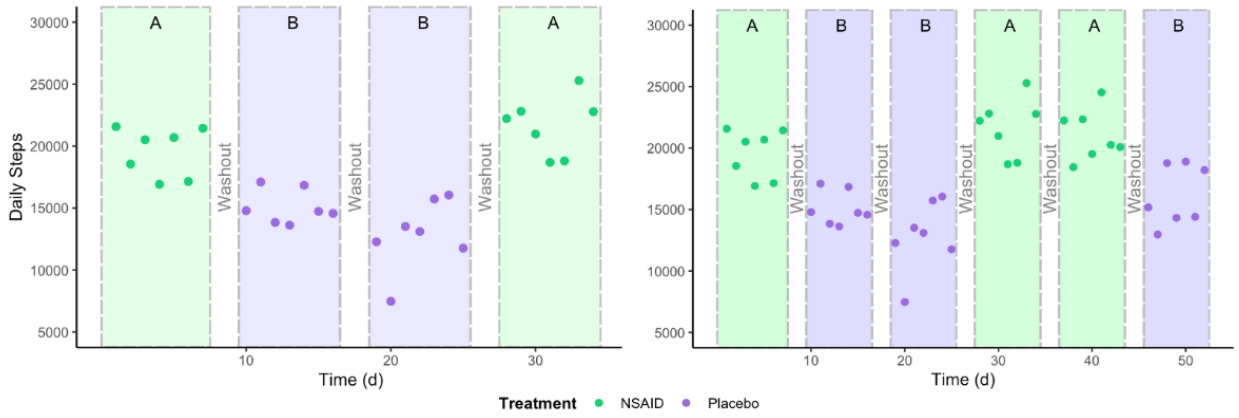

Figure 3. Hypothetical N-of-1 trials with simulated daily step count data. The effect of oral NSAID administration on daily step count is modeled in an osteoarthritis-affected thylacine (*Thylacinus cynocephalus*) using 4 (left) and 6 (right) crossover periods. Data points were sampled from two simulated normal distributions with means of 15,000 (placebo) and 20,000 (NSAID) and a common standard deviation of 2,500.

12. Interpretation of Experimental N-of-1 Data in Zoological Medicine

Strictly speaking, interpretations of N-of-1 trials should be limited to that particular patient. Practically, however, zoological medicine clinicians frequently use N-of-1 observational data in their clinical practice so it can be expected that they will also generalize the findings of N-of-1 trials. This experimental evidence would be stronger than evidence from case reports, owing to a reduction in bias achieved with placebo controls, blinding, and crossover.

 doi:10.20944/preprints202212.0424.v1

To improve generalizability of conclusions and understand the average treatment effect for the population, it is possible to perform a systematic review and meta-analysis of N-of-1 trials, as has been done in a number of human conditions.[21] In fact, these aggregated N-of-1 trials can even match the power of a conventional RCT with far fewer participants.[23] While many different methods have been used to synthesize N-of-1 trials, there is substantial debate as to which of these methods works best under any given circumstances.[12,21] As N-of-1 trials are still rare in zoological medicine, detailed discussion of the nuances of combining them are beyond the scope of this article.[12,21]

13. Experimental N-of-1 Data Conclusions

N-of-1 trials are similar to the informal therapeutic trials that clinicians so often use to manage chronic conditions because they both seek to understand how an individual patient's condition changes in response to initiation or withdrawal of treatment. N-of-1 trials differ, however, in that they are conducted rigorously using blinding, placebo controls, and carefully measured outcomes, all of which are aimed to minimize the potential for bias. This methodological rigor makes them more challenging to perform than informal therapeutic trials or SPOTs, but it allows for far greater inferences to made about the effectiveness of a treatment. While initiatives for individualized EBM are growing in prominence in human medicine, zoo clinicians often have no choice. Because we rarely have RCTs to give us an estimate of the average treatment effect, why not use N-of-1 trials to determine the effect of a treatment in the patient before us?

14. Conclusions

While suboptimal, N-of-1 data are unavoidable in zoological medicine. We are a field defined by its breadth rather than depth. Thus, it is critical that we pursue research methods that allow us to use the data we have, however few. In this article, we have introduced systematic reviews and meta-analyses of case reports to a broader audience, describing their benefits compared to narrative reviews as well as limitations compared to randomized experimental studies. We have also described prospective N-of-1 trials as a tool for generating high quality evidence of treatment effect in individual patients. Taken together, these methods utilizing N-of-1 data offer zoo clinicians a way to create and synthesize desperately needed evidence.

15. Acknowledgments

 doi:10.20944/preprints202212.0424.v1

We would like to thank Robert Goldberg, PhD and several other colleagues for their insightful comments on an earlier draft of this manuscript. This work was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, Award Number TL1TR002546 (Cummings) and the Office of Research Infrastructure Programs, National Institutes of Health, Award Number T32OD011121 (Krucik). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH nor any listed employer.

References

1. Ahn E, Kang H. Introduction to systematic review and meta-analysis. Korean J Anesthesiol. 2018;71(2):103–112.

2. Barker TH, Migliavaca CB, Stein C, Colpani V, Falavigna M, Aromataris E, Munn Z. Conducting proportional meta-analysis in different types of systematic reviews: a guide for synthesisers of evidence. BMC Med Res Methodol. 2021;21(1):189.

3. Calvache JA, Vera-Montoya M, Ordoñez D, Hernandez AV, Altman D, Moher D. Completeness of reporting of case reports in high-impact medical journals. Eur J Clin Invest. 2020;50(4):e13215.

4. Carriere KC, Li Y, Mitchell G, Senior H. Methodological considerations of N-of-1 trials. In: Nikles J, Mitchell G (eds.). The Essential Guide to N-of-1 Trials in Health. Dordrecht: Springer Netherlands; 2015. p. 67–80.

5. Cummings CO, Eisenbarth JM. Snakebite envenoming in avian species: a systematic scoping review and practitioner experience survey. J Avian Med Surg. In Press.

 doi:10.20944/preprints202212.0424.v1

6. Cummings CO, Mitchell MA, Perry SM, Fleissner N, Mayer J, Lennox AM, Johnson-Delaney CA. Bayesian decision analysis: an underutilized tool in veterinary medicine. Animals. 2022;12(23):3414.

7. Davidson KW, Silverstein M, Cheung K, Paluch RA, Epstein LH. Personalized (N-of-1) trials: a primer. JAMA Pediatr. 2021;175(4):404.

8. Effective Health Care Program, Agency for Healthcare Research and Quality. Introduction to N-of-1 Trials: Indications and Barriers (Chapter 1). Rockville, MD. 2014.

9. Effective Health Care Program, Agency for Healthcare Research and Quality. Statistical Design and Analytic Considerations for N-of-1 Trials (Chapter 4). Rockville, MD. 2014.

10. Gagnier JJ, Kienle G, Altman DG, Moher D, Sox H, Riley D. The CARE guidelines: consensus-based clinical case reporting guideline development. Glob Adv Health Med. 2013;2(5):38–43.

11. Garfield E. 100 citation classics from the Journal of the American Medical Association. 1987;257(1):52–59.

12. Hendrickson RC, Thomas RG, Schork NJ, Raskind MA. Optimizing aggregated N-of-1 trial designs for predictive biomarker validation: statistical methods and theoretical findings. Front Digit Health.2020;2:13.

13. Hopewell E, Harrison SH, Posey R, Duke EG, Troan B, Harrison T. Analysis of published amphibian neoplasia case reports. J Herpetol Med Surg. 2020;30(3):148–155.

14. Hunter RP, Isaza R. Polypharmacy in zoological medicine. Pharmaceutics. 2017;9(1):10.

 doi:10.20944/preprints202212.0424.v1

15. Journal of Zoo and Wildlife Medicine. Journal of Zoo and Wildlife Medicine Editorial Policy and Process December, 2021. 2021;51(4):1316–1326.

16. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? Per Med. 2011;8(2):161–173.

17. Maraj R, Rerkpattanapipat P, Jacobs LE, Makornwattana P, Kotler MN. Meta- analysis of 143 reported cases of aortic intramural hematoma. Am J Cardiol. 2000;86(6):664–668.

18. Margolis A, Giuliano C. Making the switch: From case studies to N-of-1 trials. Epilepsy Behav Rep. 2019;12:100336.

19. Masic I, Miokovic M, Muhamedagic B. Evidence based medicine - new approaches and challenges. Acta Inform Med. 2008;16(4): 219–225.

20. Mengersen K, McGree JM, Schmid CH. Statistical analysis of N-of-1 trials. In: Nikles J, Mitchell G (eds.). The Essential Guide to N-of-1 Trials in Health. Dordrecht: Springer Netherlands; 2015. p. 135–153.

21. Mengersen K, McGree JM, Schmid CH. Systematic review and meta-analysis using N-of-1 trials. In: Nikles J, Mitchell G (eds.). The Essential Guide to N-of-1 Trials in Health. Dordrecht: Springer Netherlands; 2015. p. 211–231.

22. Mirza R, Punja S, Vohra S, Guyatt G. The history and development of N-of-1 trials. J R Soc Med. 2017;110(8):330–340.

23. Mitchell G. Aggregated N-of-1 trials. In: Nikles J, Mitchell G (eds.). The Essential Guide to N-of-1 Trials in Health. Dordrecht: Springer Netherlands; 2015. p. 57–66.

24. Muñana KR, Zhang D, Patterson EE. Placebo effect in canine epilepsy trials. J Vet Int Med. 2010;24(1):166–170.

25. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. BMJ Evid Based Med. 2016;21(4):125–127.

26. Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. BMJ Evid Based Med. 2018;23(2):60–63.

27. Nakamura T, Igarashi H, Ito T, Jensen RT. Important of case-reports/series, in rare diseases: Using neuroendocrine tumors as an example. World J Clin Cases. 2014;2(11):608–613.

28. Newman DB, Fidahussein SS, Kashiwagi DT, Kennel KA, Kashani KB, Wang Z, Altayar O, Murad MH. Reversible cardiac dysfunction associated with hypocalcemia: a systematic review and meta-analysis of individual patient data. Heart Fail Rev. 2014;19(2):199–205.

29. Nissen T, Wynn R. The clinical case report: a review of its merits and limitations. BMC Res Notes. 2014;7:264.

30. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

31. Parkinson LA, Mans C. Effects of furosemide administration to water-deprived inland bearded dragons (*Pogona vitticeps*). Am J Vet Res. 2018;79(11):1204–1208.

32. Percha B, Baskerville EB, Johnson M, Dudley JT, Zimmerman N. Designing robust N-of-1 studies for precision medicine: simulation study and design recommendations. J Med Internet Res. 2019;21(4):e12641.

33. Pizzi R. Analysis of applications and outcomes of minimally invasive surgical techniques in wildlife. PhD Dissertation. Universitat Autònoma de Barcelona, Barcelona, Spain. 2017.

34. Rathore SS, Rojas GA, Sondhi M, Pothuru S, Pydi R, Kancherla N, Singh R, Ahmed NK, Shah J, Tousif S, Baloch UT, Wen Q. Myocarditis associated with Covid-19 disease: A systematic review of published case reports and case series. Int J Clin Pract. 2021;75(11):e14470.

35. Rochon J. A statistical model for the "N-of-1" study. J Clin Epidemiol; 1990;43(5):499–508.

36. Sampayo-Cordero M, Miguel-Huguet B, Malfettone A, Pérez-García JM, Llombart- Cussac A, Cortés J, Pardo A, Pérez-López J. The value of case reports in systematic reviews from rare diseases. the example of enzyme replacement therapy (ERT) in patients with mucopolysaccharidosis type II (MPS-II). Int J Environ Res Public Health. 2020;17(18):6590.

37. Sampayo-Cordero M, Miguel-Huguet B, Pardo-Mateos A, Malfettone A, Pérez-García J, Llombart-Cussac A, Cortés J, Moltó-Abad M, Muñoz-Delgado C, Pérez-Quintana M, Pérez-López J. Agreement between results of meta-analyses from case reports and clinical studies, regarding efficacy and safety of idursulfase therapy in patients with mucopolysaccharidosis type II (MPS-II). a new tool for evidence-based medicine in rare diseases. Orphanet J Rare Dis. 2019;14(1):230.

38. Schork NJ. Personalized medicine: time for one-person trials. Nature. 2015;520(7549):609–611.

39. Senior H. Randomization, allocation concealment, and blinding. In: Nikles J, Mitchell G (eds.). The Essential Guide to N-of-1 Trials in Health. Dordrecht: Springer Netherlands; 2015. p. 81–91.

 doi:10.20944/preprints202212.0424.v1

40. Shamseer L, Sampson M, Bukutu C, Schmid CH, Nikles J, Tate R, Johnston BC, Zucker D, Shadish WR, Kravitz R, Guyatt G, Altman DG, Moher D, Vohra S, and the CENT group. CONSORT extension for reporting N-of-1 trials (CENT) 2015: explanation and elaboration. BMJ. 2015;350:h1793.

41. Smith J, Yelland M, Del Mar C. Single patient open trials (SPOTs). In: Nikles J, Mitchell G (eds.). The Essential Guide to N-of-1 Trials in Health. Dordrecht: Springer Netherlands; 2015. p. 195–209.

42. Strong VJ, Grindlay D, Redrobe S, Cobb M, White K. A systematic review of the literature relating to captive great ape morbidity and mortality. J Zoo Wildl Med. 2016;47(3):697–710.

43. Stunnenberg BC, Deinum J, Nijenhuis T, Huysmans F, van der Wilt GJ, van Engelen BGM, van Agt F. N-of-1 trials: evidence-based clinical care or medical research that requires IRB approval? A practical flowchart based on an ethical framework. Healthcare. 2020;8(1):49.

44. Vandeweerd V. Le N of 1 trial : utile et applicable en médecine vétérinaire? Master's Thesis. Université de Liège, Liège, Belgique; 2021.

45. Whalley B, Hyland ME. Placebo by proxy: the effect of parents' beliefs on therapy for children's temper tantrums. J Behav Med. 2013;36(4):341–346.

46. Wong K, Farooq Alam Shah MU, Khurshid M, Ullah I, Tahir MJ, Yousaf Z. COVID- 19 associated vasculitis: a systematic review of case reports and case series. Ann Med Surg. 2022;74:103249.