HOW TO DESIGN AND EVALUATE SCIENTIFIC STUDIES

Ralf Mueller

Medizinische Tierklinik, Universität München, Germany

Residents in veterinary dermatology all over the world are obliged to conduct at least one scientific study in veterinary dermatology during their residency. To design such as study in a meaningful and scientific way is not easy and all too often an unexperienced resident feels overwhelmed with the task. Guidance by your mentor is essential, but not all supervisors of dermatology residencies are experienced researchers themselves. This lecture will attempt to provide some guidelines on how to conduct a study. Many of these guidelines can also be applied in the evaluation of published studies for journal club or own interest. Critical evaluation of the literature is a crucial skill for any specialist!
First, I will discuss some terms commonly used in research and then go through the process of selecting a study, conducting it, evaluating the results and writing up and submitting the manuscript. And finally I will briefly cover, which information can be also used to evaluate other studies in the literature.

Selected study terminology

Randomization
The process of randomization is aimed at the generation of treatment groups comparable in regards to known or unknown confounding factors. The efficacy of randomization depends upon the generation of allocation sequences by truly random methods and upon the adequate concealment of treatment sequences. Lack of appropriate randomization schemes could lead to *selection bias*. If you have only two groups in your study, the easiest randomization method is tossing a coin. Throwing dice is also an option with two, three or six groups in your study. A better and still readily available option is using randomization tables found in many statistical textbooks. For more involved randomization schemes a statistician should be contacted. Of course, in blinded studies, this randomization should be concealed from the researcher.

Concealment
Blinding of observers (e.g. clinicians) and participants (e.g. owners) to the treatment allocation is an essential part to avoid *detection* bias. If only clinicians are unaware of the grouping, the study is blinded, if owners are also unaware, the study is double-blinded.

Prospective study
In a prospective study, the question to be answered is posed first, the methods are planned in detail and only then is the study begun. This allows gathering of all relevant data in a thorough way. It also allows blinding and randomization and is thus the gold standard for scientific studies.

Retrospective Study
In a retrospective study, an attempt is made to answer a posed question with data gathered prior. In some instances, blinding may be possible. A good example are histopathological studies, where slides can be evaluated for certain criteria by one examiner without knowing what group

the slide belongs to. A control group may also be possible in certain retrospective studies. However, in most retrospective studies neither control groups, nor randomization, nor concealment is performed, limiting these studies significantly. These studies are often performed to strengthen anecdotal reports and should be followed by prospective studies.

Power of a study

The power of a study is the probability of a difference being detected with a fixed number of data sets. In other words, if the power of a test is 1.0 or 100%, we will always detect a difference between the groups (if it is present) and an increase in data sets will not increase our chance to find a difference. Typically, we would like the power to be at least 0.8 or 80%. As a general rule, the bigger the sample size, the smaller a difference is easily detectable. With most resident projects, the numbers are smaller and the difference between two groups or time points needs to be fairly prominent to be detected easily.

P-value

If we find a difference between two groups of data, the P-value is the probability of that difference being due to sample selection and not due to a real difference. A P-value of 0.05 (used in most studies) means that if we detected a significant difference between the groups, there is only a 5% chance that this difference was due to sampling (and will very likely not be present in the next several studies with identical protocol) and a 95% chance that we can repeat this result with the next identical study.

Significant results

The significance level is determined by the researcher. If it is stated, that a P value of 0.05 is considered significant, than any P-value ≤ 0.05 indicates significant results. 0.05 is commonly chosen, but there may be reasons to change the P-value. If you set the significance level very low (P=0.001), there is a higher chance that there is a difference, but you didn't identify it *(Type II error)*. At the same time there is a lower chance that you mistakenly identify a difference, even though there is none *(Type I error)*. This would be sensible for example in a study evaluating a drug with possible severe side effects. You don't want to have results showing a significant improvement of the disease with the drug, it gets used widely, many dogs show severe side effects and then 2 years later another study shows that your results were unlucky and in reality there was no benefit with this drug.

If however you set the significance level very high (P=0.1), there is a higher chance that you identify a difference even though there is none. But there is a smaller chance that you miss a true difference between the groups. Thus, if you examine a drug that has no adverse effects and is inexpensive, you want to minimize the chance of missing a true effect due to the possible benefits. You are less concerned about identifying a difference although there is none and dispensing this drug a lot before your results are proven wrong.

If you set your significance level higher than 0.05, you will need to explain why in the discussion of the paper!

Designing your own study

1. **Pose a question!** At the beginning of each study there is an unanswered question. Your aim is to answer that question using scientific methodology.

    - When posing a question, you should ask yourself, if you *really would like to know* the answer. It is much easier to maintain your motivation all the way to the submission of the manuscript, if in your heart you are deeply interested in your work! If the study is only performed to fulfil criteria for the board examination, may be you should search for something else.
    - It is also useful to examine, if the *answer is relevant* to the wider veterinary dermatology community. The more people are interested in your subject, the more relevant the study will be!
    - Make sure that the *question is very specific*. The more specific your question, the more likely you will get a meaningful answer.

2. **Make yourself knowledgeable!** You want to make sure, that your question has not been answered conclusively by somebody else already.
    - *Talk to your mentor and other dermatologists* about your potential study and inquire if they know about any similar studies (published or unpublished). Sometimes a study has been presented on an annual meeting, but never been published in a refereed journal. The abstract in the proceedings may not be quoted in textbooks or found in electronic searches, but may give you some more information and allow you to pose your question more specifically.
    - Perform a *thorough literature search*. Bibliographies of major textbooks in human or veterinary dermatology are good starting points. Electronic searches are an invaluable tool, they cast the net wide and often bring unexpected results. When performing this electronic search, it may be useful to have the assistance of a librarian and to look in several databases. Medline is a human classic, but CABabstracts, Agricola and others are covering journals relevant for veterinary medicine, that may be missed with Medline or other databases aimed at human medical literature. Once you obtained the articles identified by your search, evaluate the bibliography of each identified article for further relevant information.
    - If your intended question has been the topic of a published article already, evaluate the article to identify, if the study was performed properly and scientifically. If it answered your question conclusively and to your full satisfaction, you may need to find another research project. However, often a single study may not answer a question without doubt, as many times scientific studies are characterized by a comparatively small number of patients and corroboration by a second study establishes more useful evidence. Other times, you may not agree with the methods used. Or you may not agree with the conclusions drawn and may want to perform your study, possibly with different methods, to establish if your assumptions reflect the reality better.

3. **Define a feasible protocol!**
    - You are in a dermatology residency and as such have a *limited time span* for conducting your project. Remember, that it is realistic to assume that it will take up to 3 months to establish a protocol, 3 months to write the manuscript and up to 6-8 months from submission to acceptance of a paper (if everything goes smooth!!). This leaves one to two years to actually conduct the study, if you start early in your residency! If you perform a clinical study with dermatology patients, you need to consider how long it will take you to gather these patients. It is safe to suggest doubling the assumed time span for patient collection!
    - Based on the above estimations, *prospective studies* with clinical patients should examine a common disease and should probably not extend more than 3 to maximal 6 months for the actual clinical part. Prospective studies are certainly the most valuable clinical research, as they allow exclusion or at least minimization of confounding factors (factors influencing your results), avoid frustrating lack of individual data that are part of most retrospective studies and allow inclusion of a control group and randomization (see below), but they are also the most time consuming! They are typically performed when retrospective studies and/or anecdotal data are available, either to corroborate the results or to examine more detail.
    - *Laboratory studies* can be performed scientifically and in a timely fashion if the necessary resources are available (see below). As a resident, it may be prudent to avoid studies, where laboratory methods have not yet been established and need to be developed for the study, as this (in conjunction with establishing normal ranges or negative and positive controls that are often essential) may be very time consuming.
    - *Retrospective studies* are quite commonly performed by residents, as they do not have to rely on patients coming in or laboratory methods being developed. They also typically are inexpensive to conduct, they often rely exclusively on time, expertise, medical records and/or histopathology specimens. But they also are the least valuable type of study, as it is by nature difficult to include a control group or to perform blinding or randomization. Retrospective studies are typically done to "cast a net wide", to examine a topic where not much information is available. Due to the retrospective nature, they tend to ask less specific questions and their results are often the basis for future prospective studies.

4. **Locate possible resources!**
    - *Intellectual resources* are a very important part of designing a study. Is your mentor or another person at your institution an expert in the field you want to examine? Is this person easily available? Is she/he willing to help and give you the guidance and input that you will need? If not, is there a person outside your institution that has that expertise? In the electronic or information age, intellectual input can be obtained quickly and reliably from anywhere in the world via email, although a contact person at your institution would be preferable. But make sure you give your collaborator(s) a detailed idea of what you plan to do to allow them to estimate the time involved and give you optimal input!
    - *Financial resources* may be needed to a variable degree depending on your study. Do you have a resident fund that can cover these expenses easily and quickly? If

not, does your mentor have access to some funds that she/he is willing to dedicate to your study?

- Are the *materials*, instruments, test kits, laboratory tests and/or drugs needed for the study easily available at your institution? If not, where and how can you get them? How long does it take to obtain what you need? How much does it cost? Do you have to send samples? Are there special shipping requirements? How reliably can you get the results back quickly?!?

5. **Write your section about "Materials and Methods"!**
    - This section should describe what you plan to do in enough detail for other people to duplicate your research. If certain parts such as complicated laboratory tests were reported in detail previously, an abbreviated reference following "… as reported elsewhere (Reference)" is usually acceptable.
    - The *overall design* should be stated (i.e. "placebo-controlled, double blinded, randomized study" or "retrospective study")
    - The *subjects should be clearly defined*. Inclusion criteria and exclusion criteria should be defined.
    - *Treatment groups* should be decided. Ideally, a double-blinded, placebo controlled study is conducted. However, it would be unethical to have a placebo group in a study of severe skin diseases. Animal Care and Use Committees in most institution will not allow such protocols and owners will be reluctant to participate. In any severe disease or any study of longer duration, it may be more sensible to have the standard therapy as a control group.
    - *Endpoints of the study* must be determined. Do we exclusively look at clinical improvement? Then we need to develop a clinical scoring system (or adopt an existing one). Or do we want objective parameters? And which ones are most appropriate?
    - The *exact medications and doses* need to be determined as well as which other drugs are permitted during the study and which are not. We can allow all medications not assumed to potentially interfere with our medications or our disease, but need to specify exactly which ones are allowed and which ones are not. It usually is better from a scientific point of view to limit the concurrent medication as much as possible! In some studies withdrawal times from medications prior to inclusion also need to be specified.
    - The *duration of the study* and the *time points to evaluate patients* need to be determined. For a resident's study, realistically one should limit the study duration to 3 or at most 6 months. That of course does not allow the follow-up required in many clinical studies, which may not make the study useless but needs to be recognized as a potentially important limitation and addressed in the discussion. The number of revisits in clinical studies should be as small as possible to increase owner compliance.
    - If there is more than one person involved in a clinical study, the same patient preferably should be evaluated by the same clinician. In clinical studies, *detection bias* is possible due to the variation of evaluations between clinicians and time points associated. It should be minimized by clearly and specifically defining evaluation criteria.

- The *number of animals needed* can be determined statistically prior to beginning the study, when the power of the study and the P-value are decided upon and the approximate standard deviation of the data can be estimated based on previous studies. If there is no previous evidence at all allowing an educated guess on the standard deviation, then one should perform a pilot study first, which may be difficult due to the time restraints of a residency project. If enrolling patients solely based on availability is sufficient, will only become apparent at the end of the study and carries the risk of meaningless results after a lot of work and effort.
- *Statistical methods* need to be decided upon. Most scientific data can be analyzed using straightforward and simple techniques. However, if neither you nor your mentor have statistical expertise, it is prudent to seek the advice of a statistician prior to beginning the study.
- If you are employed by academic institutions, you need to submit the protocol of any study involving life animals to the *Animal Care and Use Committee* or Ethics Committee as soon as you have completed this section to allow rapid processing. It pays to emphasise the potential benefits of your study to the patients in the study as well as future patients. If only blood sampling, treatment and clinical evaluations are involved, approval is usually not a problem. If however more invasive procedures are planned, I would recommend calling a member of the Committee and asking for some input on the protocol before submitting it.

6. **Create "Cheat sheets"!**
    - However thorough a study is prepared, almost always it is conducted under certain pressures (time pressure being the most common). It is easy to overlook or forget things, particular in a clinical setting with other patients waiting, students or owners distracting etc. To facilitate the performance and minimize data loss, a sheet should be prepared describing very briefly to all clinicians and technicians involved in the study, what protocol to follow. This could be posted on the walls of all treatment and examination rooms or laboratories used.
    - For each evaluation a separate sheet should be filled out. This sheet should contain a clear identification of the animal and/or sample with name and number, the date of the evaluation, the name of the clinician performing the evaluation and all the data that needs to be gathered.
    - Cheat sheets should be easy and save time. It certainly pays to spend more time creating these well, using multiple choice answers where possible, to allow speedy completion. Double check at the end of each evaluation that the sheets are filled out completely!
    - As a general rule most researchers will find at the end of a study that they would have liked to gather more data than they actually have, as answering one question with a study often poses several others, that may possibly be or have been answered with the same or an extended data set from the same patients. It is very important to think about what other data possibly could be needed (and often easily gathered) during a study to avoid the frustration at the end!!

7. **Gather the data!**
    - How quickly patients can be enrolled in clinical prospective studies depends on two factors:

- How often will you encounter patients with the disease in question. The frequency of the disease may be increased in your practice by informing referring veterinarians about the study and asking for referrals of this disease specifically. That approach works particularly well, if owners of enrolled dogs are offered special benefits.
- How often can you enroll eligible patients. This depends on your communication skills, the type of procedure required and the effort needed to complete a study.
- Client compliance is increased, if potential benefits (refunds of consultation fees, dog food, etc.) are distributed only after the study is successfully completed!
- If a prospective study is performed, dedicate time to call owners to make sure, recheck appointments are scheduled in a timely fashion and owners will be able to come. This also increases your client compliance.
- If laboratory specimens are shipped to collaborators, make sure that these coworkers know about the shipment and expect the arrival, then call to confirm arrival. If also may pay to frequently inquire about the progress and request interim results to make sure things progress in a timely fashion.

8. **Evaluate the results!**
    - Evaluate the results using the statistical methods decided on before.
    - In scientific papers, statistical analysis is emphasised and results of your study should clearly show the statistical significance and confidence intervals of your data.
    - If there is no statistically significant evidence, don't try to create it. "A trend to significance" or "Values tended to be greater in the treatment group" are not accepted by most reviewers. If you conducted a pilot study, simply state the values and standard deviations and interpret implications in the discussion of your manuscript.

9. **Write the article!**
    - The *title* should be short and state precisely what the paper is all about.
    - The *introduction* should briefly review the literature. Don't elaborate, that is done in the discussion section of your paper. Reference each statement, preferably with a scientific article, text book references should be avoided. The exceptions are statements of textbooks that do not have their basis in a scientific study but rather in anecdotal observations, if these statements are crucial to your study. Make sure that the need for your paper is convincingly demonstrated in the introduction. This section should close with a sentence stating your aim or purpose of the study.
    - The section on *materials and methods* should be completed already prior to beginning the study and should only be adapted to the requirements of the intended journal.
    - In the *results* you present the new evidence generated to answer the question posed. It should not contain anything else. How you present these results depends on your data, but will make a huge difference in how quickly readers will grasp your results. Graphs or tables are preferred for numerical data. With small numbers of patients and unpaired measurements, a vertical column scatter plot may be considered. If the values are paired, a box and whiskers configuration with lines connecting the paired values may be more useful. If you have a larger

number of values, the above graphs become to busy and confusing and a vertical bar graph with lines for standard deviations or points with error bars may be more appropriate.

- The objective of the *discussion* is to explain the effect of your findings on the initial question. It should start with a concise summary of your main conclusions, followed by evidence of previously published papers supporting your findings and conclusion. If other evidence is contradicting your findings (or your conclusions), it needs to be mentioned and possible reasons for the discrepancies discussed. These reasons may include differences in study populations, assay sensitivities or statistical evaluations. After weighing the evidence presented, you should draw a final conclusion and close with a statement of the effect of the study on clinical practice. If no final conclusion could be made, the final statement should suggest how a conclusion could be reached with further studies.
- *Authorship* is a very controversial topic with scientific publications. Most journals have clear criteria such as contributions to design and conduct of the study and critical reviewing of the manuscript. For many journals, technical assistance, reviewing of the manuscript and providing patients or material is not sufficient for authorship. However, in academia publishing is very important and thus many academicians will appreciate being included as a coauthor. Furthermore, technical assistance may include any assistance from showing you how to run a certain assay to spending days and in some instances weeks to gather and organize data to allow you rapid statistical evaluation. I recommend, that the criteria for authorship are discussed frankly and very early on in the study with all people involved to prevent any misunderstandings or bad feelings later on. Any person that has contributed in any way and is not coauthor should have their contribution *acknowledged* at the end of the paper. It is sensible and courteous to ask for permission from anyone whom you wish to acknowledge!

10. **Submit the manuscript for publication!**
    - The choice of the journal to submit your manuscript to is very important. There are a variety of criteria to consider.
    - Your study must be of interest to the readers of the journal. Trying to publish a study on canine demodicosis in a human gastroenterology journal is doomed for failure.
    - The quality of your study must be approximately equivalent to the quality of other publications in this particular journal. The more prestigious the journal, the higher the relevance of an acceptance, if you consider an academic carrier. The more prestigious journals however are not always read by the largest veterinary audiences and a clinically highly relevant paper may be read by many more people if it is published in a veterinary journal with a wide distribution. The citation index will tell you how commonly articles published in this journal are cited in other publications. And last, if you need the article to be accepted quickly, it may be prudent to submit it to a journal with a less stringent quality control.
    - Some journals will charge a fee for submission, independent of acceptance. Others only charge a fee, when the paper is accepted. Some journals only charge a fee, if you wish to include colour pictures. And some journals don't charge a fee at all.

- There are some journals that are prestigious, distributed and cited widely and would be perfectly suitable. But to insiders they are known for reviewers eager to change many minor details, sometimes several times, before the paper is finally accepted. Others may be known for a very slow turnaround. If you are under time pressure, because your board examination participation is dependent on acceptance of your paper, try to avoid such journals. Journals with the option of electronic submission often have a faster turnaround. Mentors or other people in the field with a wider publication experience should be able to give you some advice in this regard. Many journals state in the title page of each article, when the manuscript was submitted and when it was accepted, which also gives you an idea about the average turnaround.
- Once you decided on a journal, you need to read the instructions for authors VERY carefully (a link to these instructions is typically found on the homepage of the journal). You should make absolutely sure, that the manuscript in all parts (references, foot notes, tables, figures) complies in every detail with the instructions given to authors to avoid unnecessary delays in acceptance.
- Each author then needs to read and approve the final manuscript and in some journals also sign the letter of submission to the editor, which should accompany the manuscript. The letter must clearly identify the paper, give all your contact details and fullfill any other specific requirements the journal may have.
- If this is a publication needed for board certification and timing is an issue, mention this in the submission letter to the editor, this may possibly lead to a more rapid turnaround.

Evaluating other studies

The evaluation of other studies follows the same guidelines as designing your own study.

- Usually, you evaluate the *way, the study was conducted*, by scrutinizing the section "Materials and Methods" as described above. First, you need to examine if the methods used are appropriate for answering the question posed! It is important to identify any possible bias due to selection of patients and evaluation of study criteria. The statistical methods used should be appropriate. Then you look at the data in the "Results section". Is the data variability very large (which typically makes it more difficult to evaluate clinical relevance even in face of significant differences). Is there an overlap between data sets? Are there significant differences between data sets?
- The next point is to evaluate the conclusion drawn from the obtained results. Do the results clearly support or refute the hypothesis stated? Did the "Discussion" section (as stated above) discuss all relevant published evidence, correlate it with the obtained results and explained the findings logically? And if not, are several possibilities given for the unexpected outcome? The ultimate question at the end of your evaluation is of course: "Do you trust the conclusion of this study?"
- Lastly, you can *evaluate the way the article is written*. A study can be conducted well, all the conclusions may be valid, but the discussion may miss important points, the content may be hard to understand upon first reading and grammatical errors or unsuitable figures or tables may decrease the pleasure of reading it. A good manuscript is clear, interesting, complete and a pleasure to read.

References

International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *British Medical Journal* 1988, 296: 401-405.

Motulsky H. *Intuitive Biostatistics*. Oxford University Press 1995.

Mason IS. Writing and publishing a paper in a veterinary journal. *Journal of Small Animal Practice* 1995, 36: 214-220.

Petrie A, Watson P. *Statistics for Veterinary and Animal Science*. Blackwell Science, Oxford, 1999.