Aus dem Department für Anästhesiologie und perioperative Intensivmedizin
der Veterinärmedizinischen Universität Wien

Klinik für Kleintiere und Pferde
(Leiterin: Univ.-Prof. Dr.med.vet. Maria Paula Larenza Dipl.ECVAA)

Development and preparation of a systematic review protocol, including data extraction and evaluation forms in veterinary medicine

Diplomarbeit

Veterinärmedizinische Universität Wien

vorgelegt von
Dipl. Kfm (FH) Andreas Palluch

Wien, im April 2020

Betreuerin: Dr.med.vet. Christina Braun Dipl.ACVAA

Gutachterin: Priv.Doz. Dr.med.vet Eva Eberspächer-Schweda Dipl. ACVAA FTA

Table of content

1. Introduction

The importance and awareness of evidence-based veterinary medicine (EVM) increased rapidly over the last decade, but there is still further work needed to increase full awareness and use of EVM (Huntley et al. 2017). According to the Center of Evidence-based veterinary medicine at the University of Nottingham "[EVM] is the use of best relevant evidence in conjunction with clinical expertise to make the best possible decision about a veterinary patient" (University of Nottingham, Center of Evidence-based veterinary medicine). Given the explosion of medical literature in general, every clinician faces at least two challenges. On the one hand, it is time consuming to view all available publications. On the other hand, it is difficult to identify the best evidence within the body of literature (Gopalakrishnan and Ganeshkumar 2013). To support veterinarians in gathering and understanding the best relevant evidence and hence making the best possible decision, systematic reviews are one of the most important means. The growing number of reviews published in veterinary medicine highlights the need and acceptance for reviews. As shown in ***Fig 1*** the number of published reviews in veterinary medicine more than doubled from 597 in 2013 to 1.303 in 2018 (US National Library of Medicine National Institutes of Health).

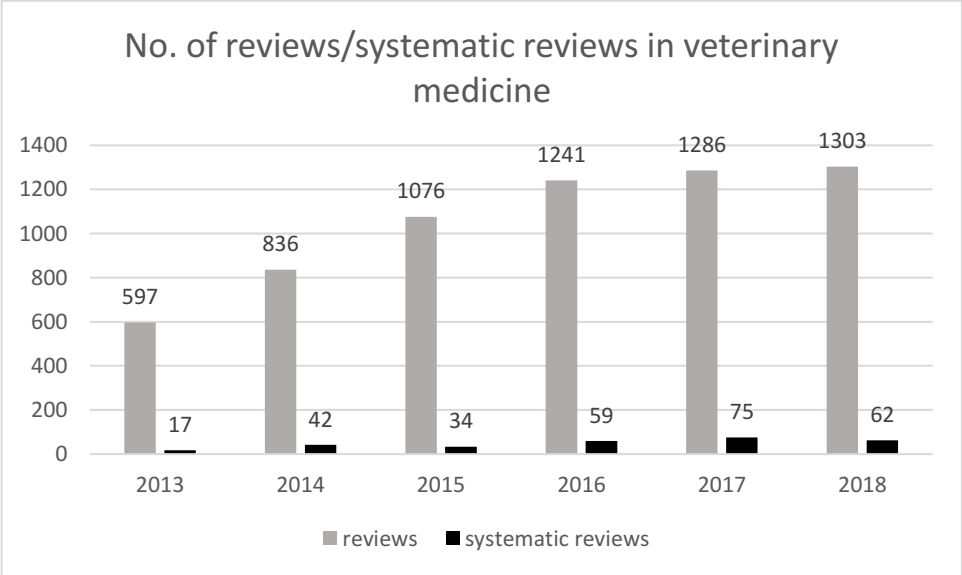


Fig 1: Number of reviews and systematic reviews in veterinary medicine listed on pubmed.org during the years 2013 to 2018. (US National Library of Medicine National Institutes of Health)

In human medicine the total number of reviews is substantially higher compared to veterinary medicine with 56791 in 2018 (search string “review human medicine” PubMed.org 17.12.2019). Of these reviews, 16.2 % (*i.e.*, 9194) are classified as ‘systematic review’ (US National Library of Medicine National Institutes of Health). The progressive increase in listed systematic reviews in general started at the beginning of this century, which matches the first development of comprehensive guidelines and frameworks. By comparison in veterinary science, much less attention has been paid to the quality of research synthesis (O'Connor et al. 2014a). Approximately five percent of all listed reviews in veterinary medicine are classified as systematic reviews (compare ***Fig 1*** 2013-2018). This is in contrast to the number of available guidelines, frameworks, and recommendations for conducting a systematic review in medicine in general and in laboratory science. That said, to the authors’ knowledge currently no specific guideline or template for clinical systematic reviews in veterinary medicine exist.

The **objective** of this thesis was to provide an overview of selected systematic review protocols, guidelines and recommended tools that could be useful for clinical systematic reviews in veterinary medicine. This will be exemplified by the development and the preparation of a **systematic review protocol** including data extraction and evaluation template for a specific review question.

2. Literature overview

One of the most known and accepted authorities on systematic reviews in medicine is the Cochrane Collaboration founded in 1993 (Cochrane). The handbook clearly defines the process step by step. Many of these steps require peer review before proceeding to the next one, which ensures a high level of evidence, reliability and transparency. These characteristics are considered the foundation for any systematic review. Indeed some of the guidelines used by the Cochrane Collaboration can be found in other systematic review protocols as well. An example is the **P**referred **R**eporting **I**tems for **S**ystematic Review and **M**eta-**A**nalysis **P**rotocols (PRISMA-P), which provides a checklist for each item that is supposed to be included in a systematic review protocol (Shamseer et al. 2015). The Cochrane Collaboration includes only systematic reviews of studies in human medicine.

More recently, research organizations started to focus on systematic reviews about animal studies and developed specific recommendations and templates. Examples of such research organizations include the **C**ollaborative **A**pproach to **M**eta-**A**nalysis and **R**eview of **A**nimal **D**ata from **E**xperimental **S**tudies (CAMARADES) and the **SY**stematic **R**eview **C**enter for **L**aboratory animal **E**xperimentation (SYRCLE). The latter center published an article about how to structure a protocol for systematic reviews of animal interventional studies (Vries et al. 2015). CAMARADES on the other hand focuses on consulting, supervising, and publishing systematic reviews and systematic review protocols on animal studies.

While Cochrane focuses only on systematic reviews about human medicine, CAMRADES and SYRCLE extend existing guidelines on animal studies. O'Connor and Sargeant (2015) were the first and so far only authors to describe systematic reviews for veterinary science. Still lacking, however, are comprehensive guidelines or detailed instructions for conducting a systematic review.

Proper evaluation of evidence is one of the most important tasks of a systematic review. Foremost is the assessment of **bias**, which can be regarded as the main challenge when evaluating evidence. Risk of bias is defined as the risk of "a systematic error or deviation from the truth, in results or inferences" (Higgins 2008). This is the exact inverse of internal validity, which is defined as "the extent to which the design and conduct of a study are likely to have prevented bias" (Higgins 2008) or "the extent to which the results of a study are correct for the circumstances being studied" (Jüni et al. 2001). The Cochrane Collaboration assesses the bias

through five domains: selection, performance, attrition, reporting, and other (Higgins 2008). Other working groups such as **G**rading of **R**ecommendations **A**ssessment, **D**evelopment and **E**valuation (GRADE) assess the risk of bias within an additional domain, *i.e.* publication bias (Guyatt et al. 2011a). This includes biases that arise through studies that remain unpublished, for example due to undesirable results. Identifying these biases requires personal expertise, experience in the field, and depth of knowledge. Thus, the demands on researchers conducting systematic reviews go beyond solely reading one or even a few studies about a given research question to assess publication bias properly.

Conducting a systematic review requires planning and multiple steps. Careful preparation and documentation of each step requires diverse skills and knowledge. For example, a librarian might be very skilled in literature search, but could lack the background information or expertise possessed by an experienced researcher to assess the quality of a study. Therefore, an entire **team** of professionals should work together to produce a systematic review (The University of Toledo).

A meta-analysis can be added to a systematic review to quantitatively analyze a body of research. "Meta-analysis is a quantitative, formal, epidemiological study design used to systematically assess the results of previous research to derive conclusions about that body of research.[…] Outcomes from a meta-analysis may include a more precise estimate of the effect of treatment or risk factor for disease, or other outcomes, than any individual study contributing to the pooled analysis." (Haidich 2010). The goal of a meta-analysis in an animal systematic review is typically to assess the general direction and magnitude of the effect of an intervention and to further explore potential sources of heterogeneity (Vesterinen et al. 2014).

Moher et al. (2015) state that "systematic reviews should build on a protocol that describes the rationale, hypothesis, and planned methods of the review". According to this definition, a review protocol is the foundation of a successful and well-performed systematic review and is therefore essential for every systematic review. In recent years, new protocols and guidelines were developed to improve the quality of systematic reviews.

In the following two subsections, methods used by Cochrane and adaptions for veterinary science are presented (Higgins 2008, Hooijmans et al. 2014). The first subsection focuses on protocols or guidelines, which specify general approaches for conducting a systematic review protocol, *e.g.* PRISMA-P and SYRCLE. The second subsection describes tools that are

available to implement specific parts of the protocol like PICO (Problem/Patient/Population, Intervention, Comparison, Outcomes), GRADE and CONSORT (CONsolidated Standards Of Reporting Trials).

2.1. Protocol templates/guidelines

PRISMA-P

***P**referred **R**eporting **I**tems for **S**ystematic Review and **M**eta-**A**nalysis **P**rotocols*

In 2015, an international group of experts created a guideline to improve the transparency, accuracy, completeness and frequency of documented systematic review and meta-analysis protocols (Shamseer et al. 2015). This guideline contains a checklist with 17 items as minimum components of a systematic review protocol. ***Tab 1*** illustrates the checklist and a short description for each item. Shamseer et al. (2015) elaborates and explains more about the checklist topics in their publication.

Tab 1 PRISMA-P checklist

Section and topic	Item No.	Checklist item
Administrative Information		
Title:		
Identification	1a	Identify the report as a protocol of a systematic review
Update	1b	If the protocol is for an update of a previous systematic review, identify as such
Registration	2	If registered, provide the name of the registry (such as **PROSPERO**) and registration number
Authors:		
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments
Support:		
Sources	5a	Indicate sources of financial or other support for the review
Sponsor	5b	Provide name for the review funder and/or sponsor
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol
Introduction		
Rationale	6	Describe the rationale for the review in the context of what is already known
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (**PICO**)
Methods		

Section and topic	Item No.	Checklist item
Eligibility criteria	8	Specify the study characteristics (such as **PICO**, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated
Study records:		
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators
Data items	12	List and define all variables for which data will be sought (such as **PICO** items, funding sources), any pre-planned data assumptions and simplifications
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I2, Kendall's τ)
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as **GRADE**)

Tab 1: PRISMA-P checklist (Shamseer et al. 2015).
PROSPERO: International Prospective Register of Systematic Reviews; PICO: Population, Intervention, Comparators and Outcomes of interest; GRADE: Grading of Recommendations Assessment Development and Evaluation

SYRCLE

SY*stematic* ***R****eview* ***C****enter for* ***L****aboratory animal* ***E****xperimentation*

According to PRISMA, veterinarian scientists have recognised that pre-specifying the **review methodology** in a protocol is an important method to promote high-quality systematic reviews (Vries et al. 2015). In contrast to human clinical systematic reviews, there was until recently no standard protocol format for systematic reviews of animal studies available.

The SYRCLE research team was one of the first to present a template of a protocol tailored to the preparation, registration and publication of systematic reviews of animal intervention studies. Previously, some authors of systematic reviews of animal studies used adapted protocols from systematic reviews of human studies (as found *e.g.* on www.dcn.ed.ac.uk /camarades/research.html#protocols). These protocols however vary in the type and amount of information included (Vries et al. 2015). The template developed by SYRCLE is based on the Cochrane review protocol and the PRISMA-P checklist and contains 50 items. These 50 items have been consolidated within the International Prospective Register Of Systematic Reviews (PROSPERO), which is managed by the National Institute for Health Research (NHS, University of York, UK). This consolidation facilitated the registration for animal studies, which follows a recommendation from all experts on systematic reviews (Vries et al. 2015).

2.2. Tools

PICO

P*roblem/Patient/Population,* ***I****ntervention,* ***C****omparison,* ***O****utcomes*

As with any research, the first step in preparing a systematic review is determining the **focus** of the study (Higgins 2008). Richardson et al. (1995) described the parameters necessary for a well-built clinical question as the key to evidence-based decisions in human medicine. The elements of the PICO question includes “P” for problem or patient or population, “I” for intervention or exposure, “C” for comparison and “O” for outcomes. This was developed to help guide a standardized and disciplined way of **formulating a clinical research question**, carrying out a thorough literature search to answer that question, and subsequently addressing an evidence-based answer to the constructed clinical query. One of the main objectives of the PICO framework is to establish a structured, reproducible, and comprehensive search strategy

(Leenaars et al. 2012). It is recommended to phrase the PICO question into appropriate search strings to gather all the relevant literature available within a given database (Saaiq and Ashraf 2017, van Loveren and Aartman 2007).

An example of an appropriate search string is (Murdoch University):

[1] Phantom limb pain OR Postamputation pain OR Post amputation pain AND

[2] Men OR Man OR Male AND

[3] Amputee OR Amputation AND

[4] Gabapentin AND

[5] Placebo AND

[6] Pain symptoms OR Neuropathic pain symptoms

Line 1, 2 and 3 represent the 'problem/patient/population', line 4 represents the 'intervention', line 5 the 'comparison' and line 6 the 'outcome'. In this case, the outcome was defined as a decrease in pain symptoms after intervention.

The SYRCLE team has adapted and improved the PICO mnemonic for application to systematic reviews of animal studies (Leenaars et al. 2012).

GRADE

***G**rading of **R**ecommendations **A**ssessment, **D**evelopment and **E**valuation*

After formulating a clinical question, it is of great importance to specify the **outcomes of interest** in as much detail as possible. The GRADE guidelines provide a tool that allows authors of systematic reviews to **rate the quality of evidence for each outcome of interest** and further helps to present the results in a tabular format. According to the GRADE working group, many if not most systematic reviews fail to address key outcomes associated with an intervention (Guyatt et al. 2011b). The GRADE working group recommends defining primary and secondary outcomes during the protocol phase to avoid reporting bias. The group also provides a classification that distinguishes outcomes that are "limited importance" or "important" from those of "critical" to the patients. This classification is comprised of a numerical rating system on a 1-9 scale (1-3, limited importance; 4-6, important; 7-9, critical).

The GRADE working group published 20 guidelines to describe the process of grading the level of evidence and how to establish summary of finding tables. Ackley (2008) defines the level of

evidence as a grade to a study, which is based on methodological quality of the study design, validity, and applicability to patients. To create an evidence profile and thus rate the level of evidence for each outcome of interest, GRADE identifies five criteria for downgrading and three criteria for upgrading the level of evidence. The starting point for each outcome depends on the study design. If it is a randomized controlled trial study (RCT), the rating starts with a high level of evidence, whereas observational studies always start at a low level of evidence. More details to down- and/or upgrading the level of evidence are provided in ***Tab 2***.

Tab 2 Quality assessment criteria

<table>
<tr><th>Study design</th><th>Quality of evidence</th><th>Lower if</th><th>Higher if</th></tr>
<tr><td rowspan="2">Randomized trial ➔</td><td>High</td><td rowspan="4">Risk of bias
-1 Serious
-2 Very serious

Inconsistency
-1 Serious
-2 Very serious

Indirectness
-1 Serious
-2 Very serious

Imprecision
-1 Serious
-2 Very serious

Publication bias
-1 likely
-2 Very likely</td><td rowspan="4">Large effect:
+1 Large
+2 Very large

Dose response
+1 Evidence of a gradient

All plausible residual confounding
+1 Would reduce a demonstrated effect or
+1 Would suggest a spurious effect when results show no effect</td></tr>
<tr><td>Moderate</td></tr>
<tr><td rowspan="2">Observational study ➔</td><td>Low</td></tr>
<tr><td>Very low</td></tr>
</table>

Tab 2: Quality assessment criteria.
Criteria for down- and/or upgrading outcomes of interest according to the GRADE guidelines (Guyatt et al. 2011b). Read table left to right. Based on the study design quality of evidence will be initially assigned from high to very low. After further assessment this level of evidence will remain, or will be downgraded or upgraded.

Inconsistency refers to the intervention effect and indicates whether differences in effects across studies can be explained by the definition of population, intervention, outcomes, or study design. If large variability in magnitude of an effect (*e.g.* based on wide confidence interval on a slope parameter) cannot be explained, then the assessor should consider downgrading the outcome parameter for inconsistency (Guyatt et al. 2011c).

There might be multiple ways of measuring the outcome of interest, and there may be multiple possible units for each measurement. In some cases it might not be possible to measure the outcome parameter directly, and only indirect (*i.e.*, surrogate or proxy) measurements are available in the studies. Confidence in results is usually higher if there is direct evidence. Indirect evidence on the contrary uses proxies for any of the following: population, intervention, or measurement. These two approaches have consequences for the quality of evidence, which is illustrated through a hypothetical example. Researchers 1 and 2 are interested in the effect of drug A on cardiac output in dogs. Researcher 1 measures blood pressure as a surrogate for cardiac output, because they do not have access to a cardiac monitor. Researcher 2, however, has access to such a monitor and can measure cardiac output directly. The study by Researcher 1 has a lower level of evidence compared to that of Researcher 2 (Guyatt et al. 2011d). To avoid reporting-bias in a systematic review, all measurement methods, surrogates, and differences in units must be documented. The **level of indirectness** (serious, very serious) must then be defined for each category within these attributes (see also section 3.6 of case study).

To assess the level of **imprecision,** reviewers should check whether the confidence interval (CI) overlaps the effects of interest documented. If the CI threshold is not crossed, the optimal information size (OIS), which is the cumulative number of samples across all studies, should be reviewed (Guyatt et al. 2011e).

The last criterion for downgrading the level of evidence is the **publication bias**. The GRADE working group acknowledges the difficulty in identifying a publication bias by using the terms “likely” and “very likely”. For publishing an evidence profile, *i.e.* a record that provides detailed information about the evidence assessment of the included outcomes by the review authors (Guyatt et al. 2011b), GRADE suggests using “undetected” and “strongly suspected”. Therefore GRADE advises rating down a maximum of one level for suspicion of publication bias, *e.g.* information previously considered of moderate level of evidence would be graded down to low level of evidence if publication bias is present (Guyatt et al. 2011a). Common scenarios where publication bias can be expected are in industry-funded studies, small sample size RCT and early results of novel therapies. The latter is mainly because novel therapies are typically based on the work of a small number of researchers, and pharmaceutical companies might not permit the publication of negative results. The decision for rating down for publication bias can be made by funnel plots. A funnel plot is a method to compare treatment

effect against a measure of study precision to detect bias or systematic heterogeneity (Light and Pillemer 1984).

According to the GRADE working group, evidence from observational studies is generally classified as low. Therefore, GRADE provides a mechanism for rating up the quality of evidence for these studies. This is possible if 1) a large magnitude of effect exists, 2) if there is a dose-response gradient, and/or 3) when all plausible confounders or other biases increase confidence in the estimated effect (see ***Tab 2***) (Guyatt et al. 2011f).

To ensure transparency and reproducibility, the reviewer must justify each down- or upgrading by including comments. Furthermore, any discrepancies in judging are discussed within the review team. If there remain any disagreements, the protocol states an additional process to solve the discrepancies for each step. This usually requires the involvement of a previously defined tie breaker.

CONSORT

CON*solidated* ***S****tandards* ***O****f* ***R****eporting* ***T****rials*

Within their first published statement in 1995, the CONSORT group expressed concerns about the increasing **risk of bias** due to poorly reported trials in human healthcare (Schulz et al. 2010). CONSORT was developed to improve the reporting of RCTs, to help readers understand trial design, analysis, and interpretation in order to assess the validity of the results (Moher et al. 2010). The scoring system includes a 25-item checklist that provides guidance for researchers planning a RCT.

In addition to CONSORT there are several other guidelines or tools to assess the risk of bias for RCT, including “Risk of Bias for animal intervention studies” (SYRCLE’s RoB tool), “criteria for levels of evidence” [Center for Evidence-Based Medicine (CEBM): http://cebm.com; 2011], and the “Jadad scale” (Hooijmans et al. 2014, Jadad et al. 1996). The RoB tool was originally designed to assess the risk of bias for laboratory animal studies and lacks some aspects that are considered necessary for animal RCTs in general. The remaining tools are less comprehensive and therefore may not assess the risk of bias as comprehensively as CONSORT.

2.3. Search strategy

The PICO components are used to create appropriate search strings. For systematic reviews about animal studies, SYRCLE further recommends the use of a step-by-step guide included in their protocol (Leenaars et al. 2012).

The evidence gathering approach has two components:

a. Database search: MEDLINE via PubMed, Web of Science, SCOPUS and EMBASE are some of the well-known databases used for systematic reviews (Vries et al. 2015).
b. Reference search: Bibliographies of those papers that match the inclusion criteria should be searched by hand to identify any further references that are relevant. These additional studies are subject to the same screening and selection process as the original search results (O'Connor et al. 2014b).

3. Case study

Introduction Vatinoxan hydrochloride

In veterinary medicine alpha-2-agonists have a nearly 60-year history of clinical use for sedation (Clarke and Hall 1969). The 'newer' drugs medetomidine and dexmedetomidine exist already for the past 20 years and 15 years, respectively. Similarly, the development and use of alpha-2-antagonists from the non-specific yohimbine and tolazoline to the more specific atipamezole took place. In very recent years the peripheral alpha-2-antagonist vatinoxan, formerly known as L-659066 or MK467, was developed and tested in several animal trials (Enouri et al. 2008, Raekallio et al. 2010, Vainionpää et al. 2013). The claim is that due to its peripheral rather than central action, vatinoxan alleviates most of the peripheral adverse effects of alpha-2-agonists while preserving sedation (Honkavaara et al. 2008).

3.1. Objective

The aim of this work is to prepare a **systematic review protocol** to evaluate the potential reversal of peripheral cardiovascular adverse effects of alpha-2-agonists through the use of vatinoxan hydrochloride in animals, and including data extraction and evaluation template.

3.2. Team

A review team consisting of the following participants was established, and tasks were assigned based on the individuals' skills and knowledge:

- Andreas Palluch (AP) (Veterinärmedizinische Universität Wien): design study, data extraction and analysis, meta-analysis, writing manuscript
- Dr.med.vet. Christina Braun Dipl.ACVAA (CB) (Veterinärmedizinische Universität Wien): design study, data extraction and analysis, writing manuscript, supervision
- Dr.med.vet Flavia Restitutti, PhD (FR) (St. George's University Grenada): data extraction and analysis, scientific input from the field
- Dr.rer.nat. Priv.-Doz Claus Vogel (CV) (Veterinärmedizinische Universität Wien): meta-analysis, statistics
- Hofrätin Mag.med.vet. Doris Reinitzer (DR) (Veterinärmedizinische Universität Wien): optimising search strategy

- Tamara Schütz (TS) (Veterinärmedizinische Universität Wien): data extraction and analysis
- Assistant Professor PhD Carlijn Hooijmans (CH) (SYRCLE, RadboudUMC): external consultant on meta-analysis and process

For each step of the process, we agreed on specific roles for each team member. The different steps of the systematic review and responsibilities of each team member are shown in ***Fig 2***. The core team consists of CB, FR, and AP. The remaining team members provide further expertise and manpower to the team.

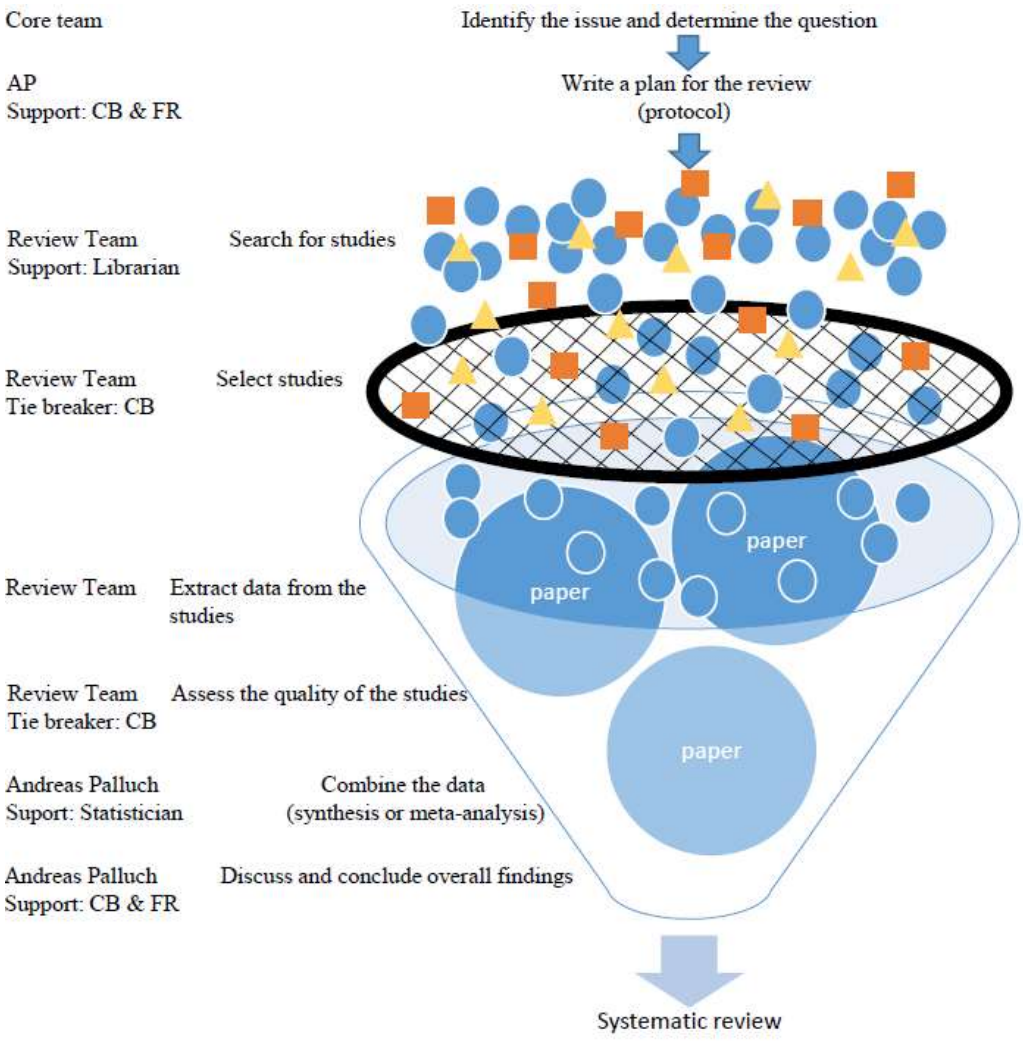


Fig 2: Systematic review process and responsibilities
AP: Andreas Palluch; CB: Dr.med.vet Christina Braun Dipl.ACVAA; FR: Dr.med.vet Flavia Restitutti

Members may enter or leave the team during the process. Additional experts such as statisticians or librarians can become involved or further reviewers might be needed. At the same time, initial members may end their participation if their skills are no longer needed or wish to be replaced by another individual.

3.3. Template selection for systematic review protocol

After an initial literature search about the subject of performing systematic reviews, all identified review protocols were assessed and discussed. The aim was to find a protocol that fits the study objective, needs minimal adjustments, and as such, should be reproducible.

The SYRCLE template (Vries et al. 2015) met those criteria and it was decided to use their template to structure the present systematic review protocol. All items were used and no adjustments were needed.

For further specification of item 38 of the systematic review protocol in 3.7 “Assessment risk of bias or study quality”, we included additional tools – GRADE and CONSORT (see 3.6).

3.4. Framing the question

Following the PICO framework the main components of the systematic review were identified including the population, intervention, comparators (or control groups), and outcomes of interest (see also section 2.2). The team began by defining the outcome parameters, as the PICO framework does not require the components to be defined in any particular order. The remaining components along with the search strategy were discussed during subsequent meetings. The study was then characterized by using the PICO framework:

- “P” for problem or patient or population: All studies regardless of animal species (mammals) or age of the patients/participants will be included. Studies of humans and non-mammals will be excluded. The rational was that due to the limited time the drug exists a limitation to a certain species at this stage would minimize the gain of information too much.
- “I” for intervention or exposure: All studies using vatinoxan to reverse (cardiovascular) effects of any alpha-2-agonists, regardless the route of administration will be included. The sole use of vatinoxan was thus explicitly not included in this study. Similarly, the

use of vatinoxan for other purposes than reversal of alpha-2-agonists was excluded to maintain a clinical focus.

- "C" for comparison: As there are no other drugs reversing only the peripheral effects of alpha-2-agonist, all studies with placebos in control groups and studies without control groups will be included.
- "O" for outcomes: At first the literature was searched for most common side effects of alpha-2-agonists used as sedatives in veterinary medicine (Murrell and Hellebrekers 2005, Sinclair 2003). Based on the obtained information 14 outcome parameters were considered. According to the Cochrane handbook a maximum number of seven outcome parameters should not be exceeded (Higgins 2008), which is in line with other recommendations and guidelines (Guyatt et al. 2011g). Thus, it was decided to reduce the number of the originally defined 14 outcome parameters to focus the review more. Out of these, four could be paired, *e.g.* "the lack of effect on alpha-2 agonist induced sedation intensity" and "the lack of effect on alpha-2 agonist induced duration of sedation" as they reflect the same field of interest. This still left twelve outcome parameters. According to Guyatt et al. (2011g), outcome parameters should be sorted according to their importance ('critical', 'important', 'limited importance'). To further eliminate outcome parameters, the twelve existing ones were ranked based on their importance. Considering clinical experience and available literature the cardiovascular adverse effects of alpha-2-agonists were considered the "critical" outcome parameters for patients (Murrell and Hellebrekers 2005, Sinclair 2003). Based on these results, the attenuation and/or prevention of cardiovascular effects of alpha-2-agonists were considered as the only (4) outcome parameters.

The research question was thus stated as: "Does vatinoxan reverse the adverse effects of alpha-2-agonists on the peripheral cardiovascular system?".

3.5. Search strategy

Consulting DR and an external process expert, CH, the following databases were searched using the predetermined strategy: MEDLINE via PubMed, SCOPUS and Web of Science.

After performing initial searches, the search strings were simplified to avoid excluding any relevant studies. Thus search results were focused only on the drug vatinoxan hydrochloride

itself and were not limited by additional qualifiers, *e.g.* by animal species. The tested search strings along with the resulting number of matches are shown in ***Tab 3***. The foundation of this literature (re)search are the inclusion and exclusion criteria (for further specifications, see points 23-29 of the systematic review protocol in 3.7). These criteria were based on the PICO approach described above (see 3.4).

According to recommendations by SYRCLE and O'Connor and Sargeant (2015), a reference search for bibliographies of those papers that match the inclusion criteria will be searched by hand to identify any further, relevant references, which will be subject to the same screening and selection process as the original search results (Leenaars et al. 2012).

Tab 3: Search strings for MEDLINE, SCOPUS and Web of Science

Search order	Database	Num. articles	Search date	Search query
1	MEDLINE	82	April 17, 2019	"L 659066"[Supplementary Concept] OR "L 659066"[All Fields] OR "l 659,066"[All Fields] OR "mk 467"[All Fields] OR "vatinoxan"[all fields]
2	SCOPUS	98	April 24, 2019	TITLE-ABS-KEY (vatinoxan OR mk-467 OR "mk 467" OR mk467 OR l659066 OR "l 659066" OR "l 659,066" OR "l659,066" OR "l-659066" OR "l-659,066") AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "re")))
3	Web of Science	98	April 24, 2019	"l-659066" OR "l-659,066" OR "mk-467" OR "vatinoxan"

Tab 3: Search strings for MEDLINE, SCOPUS and Web of Science.
Based on PICOS

3.6. Study quality assessment and data extraction

The quality assessment criteria of the GRADE guidelines were transferred into a Microsoft Excel sheet for the present systematic review and was thus included within the review protocol. It will also be used to enter all extracted information into. Then the level of evidence will be assessed by considering the results of each item according to ***Tab 2***. Starting at a “high” level

of evidence for RCTs there will be a downgrading if one or more of the criteria '**imprecision'**, '**indirectness'**, '**publication bias**', '**risk of bias**' and '**inconsistency**' are not met. The reasons for failing one of the criteria is described in section 2.2 in general and in regard to this case study below. It should be noted, that the change in grade goes by category or level. Thus, a downgrading by 2 levels of a parameter from an RCT with an initially high level of evidence would mean that now the parameter would only be considered of low level of evidence. Further downgrading would result in the lowest level of evidence (very low). More negative evaluation of the parameter will not change this result, but positive evaluation could counteract the downgrading or even improve the grading.

According to GRADE, the body of evidence for **imprecision** will be rated down if the **OIS criterion** is not met. The GRADE recommendation suggests to consider a downgrading if the sample size for a particular study is less than 400, which would lead to a downgrading in almost all cases. Another approach is to calculate the OIS and "if the total number of patients included in a systematic review is less than the number of patients generated by a conventional sample size calculation for a single adequately powered trial, consider rating down for imprecision" (Guyatt et al. 2011e). For the review process, it was decided to perform a post-hoc power analysis to calculate the optimal sample size for each outcome parameter. For repeated measurements, it was decided to calculate the power (β) for the lowest and highest effect size compared to baseline or (if available) to a placebo or control group. We agreed on an alpha of 0.05 and a correlation of 0.5. The calculation will be done by G*Power (online available at http://www.gpower.hhu.de/, Heinrich Heine University Düsseldorf). A downgrading for imprecision will be assumed if β is below 0.8.

To evaluate the **indirectness** for each outcome parameter the measurement method, the measurement unit, and if appropriate the surrogate measurement method and unit were pre-defined. If the measurement method was according to the gold standard, there will be no downgrading for indirectness. Methods that are not validated or are known to be less precise or unreliable will be graded down according to ***Tab 4***:

Tab 4 Grading indirectness

Outcome parameter	Units	Measurement method	Alternative method (impact on indirectness 0, -1, -2)	Surrogate method (impact on indirectness 0, -1, -2)
Attenuation of bradycardia	Beats/min	Pulse palpation, pulse rate measuring	Auscultation (0), ECG (0)	NA
Maintenance of cardiac output (CO)	L/min, mL/min	Ultrasonic transit-time perivascular flow probes	Dye dilution (0), thermodilution (0), lithium dilution (-1 for indirectness in combination with xylazine), Doppler (-1), non-invasive modification of Fick principle (-1)	Cardiac Index (0)
Prevention of systemic vascular resistance (SVR) increase	Dynes*s/cm5 dynes*s/cm5/ BW*0.67 dynes*s/cm5/ m2	SVR = 80 x (MAP - MVP) / CO MAP measurement => see "maintenance of MAP" MVP measurement => direct: invasive and requires catheterization of the vena cava cranialis CO => see "maintenance of CO"	MAP measurement => see "maintenance of MAP" CVP measurement => no downgrading due to clinical relevance CO measurement => see "maintenance of CO"	SVRI (0) calculation: (MAP - CVP) * 79.92 / Cardiac Index
Maintenance of mean arterial blood pressure (MAP)	mmH2O, mmHg	Direct: arterial cannulation of a central or peripheral artery	MAP measurement => indirect: oscillometric monitor (-1); ultrasonic Doppler apparatus (-1)	SAP (-1), DAP (-1)

Tab 4: Grading indirectnes.
Definition of unit, measurement method, alternative and surrogate measurements, incl. impact on indirectness. BW: Body Weight; CO: Cardiac Output; CVP: Central Venous Pressure; DAP: Diastolic Arterial Pressure; ECG: ElectroCardioGraphy; MAP: Mean Arterial Pressure MVP: Mean Venous Pressure; NA: Not Available; SAP: Systolic Arterial Pressure; SVR: Systemic Vascular Resistance; SVRI: Systemic Vascular Resistance Index

The decision for rating down for **publication bias** will be made by the core team for this review by critical visual inspection of Funnel plots. If a symmetric inverted funnel shape arises, a publication bias is unlikely. Any asymmetry will be discussed and common agreement will lead to a downgrading or not.

To facilitate the assessment of possible **risk of bias** for each RCT study, information using the CONSORT checklist will be collected (Schulz et al. 2010). For this purpose the CONSORT checklist was integrated into the interactive excel sheet. All 25 items (36 including sub-items) were screened and the outcome parameters within the table individually weighted by the core team, according to their importance to the quality of evidence. The core team assessed the

importance on basis of experience, literature, and thorough discussion. Items three to 14 and item 17 were identified to be assessed for each outcome parameter individually. To estimate the risk of bias for each outcome, the weighted points will be summed and a percentage will be calculated. We defined 50 % and 70 % as thresholds for bias. That means, if a study/outcome parameter of a study scores more than 70 % there will be no downgrading for risk of bias. If the study scores between 50 and 70 % a downgrading of - 1 will be applied. If the study scores below 50 %, a downgrading of - 2 will be applied. ***Tab 5*** lists all CONSORT items and the adjustments made.

Tab 5 CONSORT checklist of information to include when reporting a randomized trial

Section/Topic	Item No.	Checklist item	Reported	HR	CO	SVR	MAP	Weight
Title and abstract	1a	Identification as a randomised trial in the title	0					1
	1b	Structured summary of trial design, methods, results, and conclusions	0					1
Introduction								
Background and objectives	2a	Scientific background and explanation of rationale	0					1
	2b	Specific objectives of hypothesis	0					1
Methods								
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	0	0	0	0	0	1.5
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	0	0	0	0	0	0.5
Participants	4a	Eligibility criteria for participants	0	0	0	0	0	0.5
	4b	Settings and locations where the data were collected	0	0	0	0	0	0.5
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	0	0	0	0	0	1.5
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	0	0	0	0	0	1.25
	6b	Any changes to trial outcomes after the trial commenced, with reasons	0	0	0	0	0	0.25
Sample size	7a	How sample size was determined	0	0	0	0	0	0.9
	7b	When applicable, explanation of any interim analyses and stopping guidelines	0	0	0	0	0	0.1

Sequence generation	8a	Method used to generate the random allocation sequence	0	0	0	0	0	0.5
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	0	0	0	0	0	0.3
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	0	0	0	0	0	0.1
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	0	0	0	0	0	0.1
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	0	0	0	0	0	0.9
	11b	If relevant, description of the similarity of interventions	0	0	0	0	0	0.1
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	0	0	0	0	0	0.9
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	0					0.1
Results								
Participant flow	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	0	0	0	0	0	0.75
	13b	For each group, losses and exclusions after randomisation, together with reasons	0	0	0	0	0	0.25
Recruitment	14a	Dates defining the periods of recruitment and follow-up	0					0
	14b	Why the trial ended or was stopped	0					0
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	0					0
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	0					0
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95 % confidence interval)	0	0	0	0	0	1
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	0					0
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	0					0

Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	0					1
Discussion								
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	0					1
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	0					0.25
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	0					0.25
Other information								
Registration	23	Registration number and name of trial registry	0					1
Protocol	24	Where the full trial protocol can be accessed, if available	0					0.5
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	0					1
Total								***21***

Tab 5: CONSORT checklist of information to include when reporting a randomized trial (Schulz et al. 2010)
We included our outcome parameter and the weighting; HR: attenuation of bradycardia, CO: maintenance of cardiac output, SVR: prevention of SVR increase, MAP: maintenance of MAP; entries should be only made into the color-coded fields (0 = not reported; 1 = reported in the study). A total score of 21 (due to the weight of each item) is possible if all 36 items are reported in a study.

During several discussions about the topic “risk of bias” it became obvious that none of the published tools, guidelines and frameworks could be applied satisfactorily. Thus, it was agreed on CONSORT including some further adjustments regarding the weighting. During the first “test runs” with the evaluation template, there was still dissatisfaction with the process due to its subjective nature. Therefore, a subjective assessment for the risk of bias was included into the evaluation template, which is labeled as **personal level of overall trustworthiness**. The levels where chosen to match the levels of the GRADE template (levels: “high”, “moderate”, “low” or “very low”) as intentionally subjective to cover the “gut feeling” of the reviewer regarding the overall quality of each study. That means, the personal level of overall trustworthiness is a purely subjective assessment, which has not been described before. It’s purpose serves to satisfy the reviewers need to voice their personal feeling about a study/outcome parameter, which can be considered a bias in itself. It is not meant to actually judge the outcome parameter in a more objective, transparent way and should not be used as such. Yet, it might help to analyze differences in assessment between reviewers and find personal bias. It will be used on an experimental stage for the review proposed.

3.7. Systematic review protocol

SYSTEMATIC REVIEW PROTOCOL FOR ANIMAL INTERVENTION STUDIES

FORMAT BY SYRCLE (WWW.SYRCLE.NL)

VERSION 2.0 (DECEMBER 2014)

Item #	Section/Subsection /Item	Description	Check for approval
	A. General		
1.	Title of the review	A new player in town? The peripheral alpha-2-antagonist vatinoxan: a systematic review	
2.	Authors (names, affiliations, contributions)	- Andreas Palluch (AP) (Veterinärmedizinische Universität Wien): design study, data extraction and analysis, meta-analysis, writing manuscript - Christina Braun (CB) (Veterinärmedizinische Universität Wien): design study, data extraction and analysis, writing manuscript, supervision - Claus Vogel (CV) (Veterinärmedizinische Universität Wien): meta-analysis - Doris Reinitzer (DR) (Veterinärmedizinische Universität Wien): optimising search strategy - Flavia Restitutti (FR) (St. George's University Grenada): data extraction and analysis, scientific input from the field - Tamara Schütz (TS) (Veterinärmedizinische Universität Wien): data extraction and analysis	
3.	Other contributors (names, affiliations, contributions)	Carlijn R Hooijmans (CH) (SYRCLE, RadboudUMC): meta-analysis and process consulting	
4.	Contact person + e-mail address	Andreas Palluch: 1545195@students.vetmeduni.ac.at	
5.	Funding sources/sponsors	Funded by department (Veterinärmedizinische Universität Wien)	

6.	Conflicts of interest	FR is one of the co-owners of the United States Patent and Trademark Office patent application number 15/525,382, entitled “Compositions comprising substituted benzofuroquinolizine and alpha2-adrenergic agonists”. The same patent is pending under the number 2015858374 at the European Patent Office.	
7.	Date and location of protocol registration	July 16, 2019 PROSPERO	
8.	Registration number (if applicable)	142692	
9.	Stage of review at time of registration	Protocol stage	
	B. Objectives		
	Background		
10.	What is already known about this disease/model/ intervention? Why is it important to do this review?	In veterinary medicine alpha-2-agonist have a nearly 60 year old history of clinical use for sedation (Clarke and Hall 1969), incl. for the past 20 years medetomidine, the racemic mixture of dexmedetomidine and finally dexmedetomidine itself for the past 15 years. Similarly the development and use of alpha-2-antagonists from the fairly non-specific yohimbine and tolazoline to the more specific atipamezole took place. In very recent years the peripheral alpha-2-antagonist vatinoxan hydrochloride (formerly known as L-659066 or MK467) has been developed and so far tested in several animal trials (Enouri et al. 2008, Raekallio et al. 2010, Vainionpää et al. 2013). The claim is that due to its peripheral action vatinoxan alleviates most of the peripheral adverse effects of alpha 2 agonists while preserving sedation (Honkavaara et al. 2008). The aim of this review is to systemically evaluate the potential reversal of the peripheral adverse effects of alpha-2-agonists by vatinoxan. Focussing on the most important adverse effects as bradycardia, decrease in cardiac output and increase in systemic vascular resistance, we also plan to consider the influence on sedation and analgesia and ileus – especially in horses (Zeiler 2015) – (Murrell and Hellebrekers 2005, Sinclair 2003) and try to establish an overview about expected adverse effects of vatinoxan itself. This will be done in a separate systematic review. In consequence we will comprehend if this new player in veterinay medicine will help to further	

		improve the quality and safety of alpha-2-agonist induced sedation.	
	Research question		
11.	Specify the disease/health problem of interest	Cardiovascular adverse effects of alpha-2-agonists	
12.	Specify the population/species studied	Species: animals (mammals) Condition: *in vivo* Age: all ages	
13.	Specify the intervention/ exposure	Administration of vatinoxan to animals treated with alpha-2-agonists to reverse the peripheral cardiovascular adverse effects	
14.	Specify the control population	Placebo, no control group	
15.	Specify the outcome measures	- Attenuation of bradycardia (HR) - Maintenance of cardiac output (CO) - Prevention of systemic vascular resistance (SVR) increase - Maintenance of arterial blood pressure (MAP)	
16.	State your research question (based on items 11-15)	Does vatinoxan reverse the peripheral cardiovascular adverse effects of alpha-2-agonists?	
	C. Methods		
	Search and study identification		
17.	Identify literature databases to search (*e.g.* Pubmed, Embase, Web of science)	**x** MEDLINE via PubMed **x** Web of Science **x** SCOPUS □EMBASE □Other, namely: □Specific journal(s), namely:	
18.	Define electronic search strategies	Supplementary file containing search strategy: "Search strategy vatinoxan.docx" (available at request of contact person)	
19.	Identify other sources for study identification	**x** Reference lists of included studies □Books **x** Reference lists of relevant reviews □ Conference proceedings, namely: □ Contacting authors/ organisations, namely: □ Other, namely:	
20.	Define search strategy for these other sources	Screening the reference lists for relevant titles and screening the abstracts of these relevant titles.	
	Study selection		

21.	Define screening phases (*e.g.* pre-screening based on title/abstract, full text screening, both)	Three-stage screening: - Title - Abstract - Full-text	
22.	Specify (a) the number of reviewers per screening phase and (b) how discrepancies will be resolved	All identified references will be screened independently by two reviewers in each stage. FR and AP will independently review the title and abstracts. If there are any discrepancies CB will act as tie breaker. The full text review will also be done by at least two reviewers. To optimize the individual workload and risk of bias during the full text review we developed the following approach: - Three teams: CB + FR (1), CB + AP (2), FR + AP (3) - All papers listing FR as one of the authors are automaticly asigned to Team 2 - The remaining sources will be orderd by publication date and asigned: o most recent = Team 1 o most recent – 1 = Team 2 o most recent – 2 = Team 3 o … - Any discrepancies in one of the teams will be resolved by the remaining person	
	Define all inclusion and exclusion criteria based on:		
23.	Type of study (design)	Inclusion criteria: animal intervention studies; studies will be included regardless of the methodological quality Exclusion criteria: non-intervention studies	
24.	Type of animals/population (*e.g.* age, gender, disease model)	Inclusion criteria: animals (mammals) of all species, ages and sex Exclusion criteria: humans, *ex vivo*, *in vitro*	
25.	Type of intervention (*e.g.* dosage, timing, frequency)	Inclusion criteria: all administration routes and dosages will be reviewed Exclusion criteria: none	

26.	Outcome measures	Inclusion criteria: attenuation of bradycardia (HR), maintenance of cardiac output (CO), prevention of systemic vascular resistance (SVR) increase, maintenance of arterial blood pressure (MAP) Exclusion criteria: no relevant outcome measures	
27.	Language restrictions	Inclusion criteria: English, German Exclusion criteria: other languages	
28.	Publication date restrictions	Inclusion criteria: all publication dates Exclusion criteria: none	
29.	Other	Inclusion criteria: none Exclusion criteria: none	
30.	Sort and prioritize your exclusion criteria per selection phase	Selection phase title and abstract: 1. Review 2. No full paper (abstract, comment) 3. Data published in duplicate 4. Human 5. Not *in vivo* (*e.g. ex vivo*/*in vitro*/*in sillico*) 6. No use of vatinoxan 7. No relevant outcome measure 8. Other intervention (*e.g.* no use of alpha-2-agonist) Selection phase full-text: 1. Not *in vivo* (*e.g. ex vivo*/*in vitro*/*in sillico*) 2. No use of vatinoxan 3. Human 4. No relevant outcome measure 5. Other intervention (*e.g.* no use of alpha-2-agonist)	
	Study characteristics to be extracted (for assessment of external validity, reporting quality)		
31.	Study ID (*e.g.* authors, year)	Authors, title, year, contact author e-mail	

32.	Study design characteristics (*e.g.* experimental groups, number of animals)	Number of animals in experimental and control groups, presence of control group, power calculation reported	
33.	Animal model characteristics (*e.g.* species, gender, disease induction)	Animal species, strain, age, weight, gender, condition (*e.g.* healthy, sick)	
34.	Intervention characteristics (*e.g.* intervention, timing, duration)	Anaesthesia (*e.g.* no, inhalant, TIVA, PIVA), alpha-2-agonist (*e.g.* substance, dosage, route), comparison/placebo, vatinoxan (*e.g.* dosage, route, time point of administration)	
35.	Outcome measures	Presence of any other outcome measures	
36.	Other (*e.g.* drop-outs)	Number of animals excluded from statistical analysis, reason for excluding animals	
	Assessment risk of bias (internal validity) or study quality		
37.	Specify (a) the number of reviewers assessing the risk of bias/study quality in each study and (b) how discrepancies will be resolved	a) 2 reviewers. The criteria will be independently assessed by AP, CB, FR, or TS b) Differences of opinion that cannot be resolved by discussion will be resolved by invoking a third investigator (one of the remaining reviewer)	
38.	Define criteria to assess (a) the internal validity of included studies (*e.g.* selection, performance, detection and attrition bias) and/or (b) other study quality measures (*e.g.* reporting quality, power)	□ By use of SYRCLE's Risk of Bias tool □ By use of SYRCLE's Risk of Bias tool, adapted as follows: □ By use of CAMARADES' study quality checklist □ By use of CAMARADES' study quality checklist, adapted as follows: **x** Other criteria, namely: a) CONSORT: updated CONsolidated Standards of Reporting Trials 2012 checklist (25 items including identification of a control group); items 3-14 and item 17 will be assessed for each outcome parameter separately, all other items consider the complete study design; each item was discussed and weighted by the review core team (AP, CB and FR); the raters will be trained to review and grade manuscripts; b) According to Grading of Recommendations Assessment Development and Evaluation (GRADE) guidelines (Guyatt et al. 2011b), we will furthermore assess inconcistency,	

		indirectness, imprecision and publication bias Additionally we will perform a power analysis for each outcome parameter, which will be considered for the imprecision.	
	Collection of outcome data		
39.	For each outcome measure, define the type of data to be extracted (*e.g.* continuous/ dichotomous, unit of measurement)	HR: continuous; unit: beats/min; method: pulse palpation, pulse rate measuring, auscultation, ECG CO: continuous; unit: L/min, mL/min; method: Ultrasonic transit-time perivascular flow probes, dye dilution, thermodilution, lithium dilution (-1 for indirectness in combination with xylazine), doppler (-1 for indirectness), non-invasive modification of Fick (-1 for indirectness); surrogate: cardiac index SVR: continuous; unit: dynes*s/cm^5, dynes*s/cm^5/BW*0,67, dynes*s/cm^5/m^2; method: formula, imprecision due to CO or MAP; surrogate: SVRI MAP: continuous; unit: mmH_2O, mmHg; method: direct, oscillometric (-1 for indirectness), doppler (-1 for indirectness); surrogate: SAP (-1 for indirectness), DAP (-1 for indirectness); location: central, peripher (thoracic limb, pelvic limb, tail, head)	
40.	Methods for data extraction/retrieval (*e.g.* first extraction from graphs using a digital screen ruler, then contacting authors)	1. Extraction from text and tables. 2. Contacting authors by e-mail (max. 2 attempts). 3. Extraction from graphs using digital image analysis software WebPlotDigitizer Version 4.1 (available on https://apps.automeris.io/wpd/ March 15, 2019) by two independent reviewers.	
41.	Specify (a) the number of reviewers extracting data and (b) how discrepancies will be resolved	a) Two reviewers (AP and TS) will extract all data. b) Discrepancies will be resolved by discussion.	
	Data analysis/synthesis		

42.	Specify (per outcome measure) how you are planning to combine/compare the data (*e.g.* descriptive summary, meta-analysis)	Meta-analysis; if a meta-analysis is not possible for one or more outcome parameter, a descriptive summary will be done.	
43.	Specify (per outcome measure) how it will be decided whether a meta-analysis will be performed	If the studies are sufficiently comparable (with regard to design etc.), outcome data will be pooled. Subgroup analyses will only be performed, if the overall meta-analysis contains a minimum of 4 studies.	
	If a meta-analysis seems feasible/sensible, specify (for each outcome measure):		
44.	The effect measure to be used (*e.g.* mean difference, standardized mean difference, risk ratio, odds ratio)	HR: MD CO: MD (all data converted to L/min) SVR: SMD or MD (depending on reported units and whether these can be converted to a single unit) MAP: MD (all data converted to mmHg)	
45.	The statistical model of analysis (*e.g.* random or fixed effects model)	Random effects model	
46.	The statistical methods to assess heterogeneity (*e.g.* I^2, Q)	I^2	
47.	Which study characteristics will be examined as potential source of heterogeneity (subgroup analysis)	Animal species (mouse *vs.* rat *vs.* pig etc.) Gender (male *vs.* female *vs.* mixed) Alpha-2-agonist (dexmedetomidine *vs.* xylazine *vs.* medetomidine *etc.*) Route of administration (i. v. *vs*. i. m. *vs.* s. c. *etc.*)	
48.	Any sensitivity analyses you propose to perform	Animal species Alpha-2-agonist	
49.	Other details meta-analysis (*e.g.* correction for multiple testing, correction for	Where outcomes are measured repeatedly on different time points, we will use the time point at which the measured efficacy is greatest.	

	multiple use of control group)		
50.	The method for assessment of publication bias	Critical visual inspection of Funnel plots	
Final approval by (names, affiliations):		Andreas Palluch (Veterinärmedizinische Universität Wien) Christina Braun (Veterinärmedizinische Universität Wien) Flavia Restitutti (St. George's University Grenada)	Date:

The protocol for this systematic review has been online registered and is available on https://www.crd.york.ac.uk/prospero/, registration no. CRD42019142692.

3.8. Data extraction template

In addition to the protocol, an Excel sheet for data extraction and the evaluation of the level of evidence for each outcome parameter was designed. This combines the GRADE framework and the CONSORT checklist (see ***Appendix 1***).

3.9. Planned presentation of results/outlook

The systematic review itself will include a meta-analysis if there is a sufficient number of studies from which to extract data. Additionally a systematic narrative synthesis will be conducted. Narrative synthesis has been described as '(…) an approach to the systematic review and synthesis of findings from multiple studies that relies primarily on the use of words and text to summarise and explain the findings of the synthesis' (Popay et al. 2006). Summary of findings tables will highlight the main characteristics and findings of the included studies. Additional material such as a flow chart of the entire selection process [in accordance to PRISMA (Shamseer et al. 2015), see ***Appendix 2***], protocol, data extraction form, and summary of findings table will be included in manuscript for a peer-reviewed veterinary journal.

4. Discussion

A systematic review protocol and template was developed and prepared to synthesize the literature on the reversal of the peripheral cardiovascular adverse effects of alpha-2-agonist sedatives in animals with vatinoxan hydrochloride. For this purpose, the literature was first search for templates and guidelines on systematic reviews themselves. A rigorous systematic review protocol will help to guarantee a transparent, reliable and reproducible systematic review. A clearly formulated and structured clinical question, a detailed description of the review process, and establishing a comprehensive assessment will reduce errors and uncertainties during the future project. The developed approach is therefore designed to minimize reporting bias.

Template selection

At the beginning of this project, it was decided to write a systematic review about vatinoxan hydrochloride according to the gold standard for health related systematic reviews. This decision lead directly to the Cochrane Collaboration, the leading institution for high-quality, relevant, and accessible systematic reviews (Cochrane). Unfortunately, the Cochrane Library does not include animal studies. Regardless of this, it was the aim of the study authors to apply the same rigor to the present review than proposed by the Cochrane Collaboration. This led to a further search, seeking templates for animal studies. So far, there are only a limited number of paper about systeamtic reviews for animal studies. O'Connor et al. (2014b) discuss the approach to conduct a systematic review in veterinary science. They recommend the same methodology and guidelines choosen in this protocol, but they do not provide actual guidance or templates. Furthermore they discuss the challenges regarding meta-analysis for veterinary studies (*e.g.* heterogenity, number of studies and patients) in a different publication (O'Connor and Sargeant 2015). In animal science there are two organizations conducting research on systematic reviews and meta-analysis, CAMARADES and SYRCLE. CAMARADES, however, published no guidelines or templates, but focuses on consulting, supervising and publishing systematic reviews and systematic review protocols. Whereas SYRCLE is currently the only organization, which issued a genuine template for a systeamtic review protocols and published an article about how to structure a protocol for systematic reviews of animal interventional studies (Vries et al. 2015).

The SYRCLE template was chosen for the protocol and the guidelines recommended by Cochrane and O'Connor *et al.* (2014a) (*i.e.* PICO, PRISMA, GRADE, CONSORT) will be applied to the following systematic review.

While a meta-analysis could be added to the systematic review, its preparation and discussion are outside the scope of this thesis and will not be further discussed in this context.

Main challenges

Once a template and appropriate guidelines were found, the actual process of creating the topic specific protocol was challenging. Three main topics were particularly subjects of many discussions during team meetings: 1) identifying the appropriate outcome parameters, 2) how to judge the quality of evidence, and 3) specifically how to address the risk of bias.

1) Outcome parameters

The decision process as to which outcome parameter to include versus exclude was time-intensive and involved additional literature review about the topic, consultation with other specialist and experts in the field, and (re-)reading the guidelines for outcome selection. Considering the different PICO components, outcome parameters are of crucial importance to answer the clinical question through a systematic review. Guyatt et al. (2011g) state that "many, if not most, systematic reviews fail to address some key outcomes". Based on the current experience the core team like to add, that there is a risk to include outcomes of lesser importance instead and/or lose sight of the importance of different parameters. According to the Cochrane handbook a maximum number of seven **outcome parameters** should not be exceeded (Higgins 2008). In particular the clear statement made in the above mentioned handbook limiting the number of outcome parameters helped for the present protocol to reduce the complexity of the planned review, which enhances the study focus and maintains a manageable workload. Determining the outcome parameters resulted in refining the review question as well. While a well-built clinical question should be the starting point for any systematic review (Richardson et al. 1995), one should not neglect the opportunities for refinement of the research question prior to the actual start of the systematic review.

This refinement *e.g.* led also to cluster outcomes of interest resulting in three specific review question: The first and present review (protocol) focuses on the "critical" cardiovascular

adverse effects of alpha-2-agonists reversed by vatinoxan. The second focuses on sedation, analgesia, blood gases and influence on gastrointestinal effects, all outcome parameters that are considered clinically relevant, yet of lower importance compared to the cardiovascular effects. The third focuses on pharmacokinetics of vatinoxan hydrochloride. This subdivision appears necessary not only to maintain focus, but to maintain a reasonable scope of a single systematic review, as many Journals limit the word count of systematic reviews from 3,000 to 5,000 (Sampson 2014). It is expected, that only minor changes will be necessary to adapt the current protocol for the other two subject matters.

2) Quality of evidence

Whereas many authors prefer to rate a study with only one body of evidence, an overall score should not be part of a systematic review (Guyatt et al. 2011h). This is, because each body of evidence can differ significantly from one another. For the rating of the quality of evidence for the present review, each criterion will be independently assessed without an attempt to collate and assign an overall score. While this approach appears straightforward, the actual task was more difficult to the team members than originally anticipated when a first trial run was performed. To dissociate a single outcome parameter from other parameters and the general appearance of the paper felt challenging. It demands a detailed dissection of the paper, omitting irrelevant information for the parameter in question, while still being aware of more general shortcomings, *e.g.* of the subject study, and/or the authors (see also below: risk of bias).

In addition, the terminology within the GRADE guidelines and CONSORT checklist is not necessarily intuitive. The reviewer team found that it is essential not only to read their description, but also to discuss these terms and practice examples to increase coherence (Sargeant and O'Connor 2014).

The use of confidence intervals and/or the optimal information size criterion as more objective measures are constructive tools in this process.

3) Risk of bias including publication bias

The assessment of the risk of bias was perceived as one of the most challenging tasks. This might explain, why according to one systematic review about risk of bias in studies investigating the administration of IV fluid therapy in domestic animals, most studies did not comply with current evidence standards (Muir et al. 2017). Hence, it was decided to stress this

concern and to combine two tools - GRADE and CONSORT. CONSORT is helpful to assess the studies in detail according their evidence standards, *i.e.* in the template for the proposed study CONSORT will be used to assess parts of the GRADE criteria. Both GRADE and CONSORT include comprehensive checks and recommendations to avoid or at least minimize the risk of bias.

One particular type of bias, namely publication bias, needs knowledge and understanding beyond the actual parameter and paper that is to be assessed. A systematic review of a drug in its early stage such as vatinoxan bears a high risk for publication bias and imprecision (Guyatt et al. 2011e, Guyatt et al. 2011h, Guyatt et al. 2011a). This high risk exists, because there are only a few research teams studying vatinoxan. While this helps to compare and combine study results, as the methodology is the same or at least similar among studies, for the exact same reason this increases the chance of systematic mistakes and bias. Guyatt et al. (2011b) discussed these errors as the most common ones in their guidelines. First, the teams usually have only limited access to experimental animals, which leads to the use of the same animals across studies (O'Connor et al. 2014a) and may limit inferences to the group of sampled individuals. Second, these research teams typically have contracts with pharmaceutical companies who finance the studies and therefore are interested in favourable results. This might lead to delays or selective publications especially of negative results that might harm the (future) marketing of the drug or might create negative press for the company (Guyatt et al. 2011a, Guyatt et al. 2017).

Besides the use of multiple tools, interestingly, during the trial runs, it appeared that the personal ‘feeling’ about an outcome or entire paper distracted the reviewers from a transparent and rigorous determination of said parameter. Thus, additionally to the level of evidence determined by the applied GRADE template, a subjective assessment for the risk of bias was included into the evaluation template, named ‘personal level of overall trustworthiness’. Although this is not a validated way to evaluate the risk of bias, it served the purpose of relieving the need to document this subjective judgement, which possibly includes personal bias. Potentially, it could be useful as a tiebreaker and/or to explain differences in grading. In an exploratory fashion, it is planned to compare the more objective measurement of the risk of bias and the subjective level of overall (subjective) trustworthiness.

Registration

Registration of systematic-review protocols is mandatory for all Cochrane reviews and is becoming increasingly popular among researchers. Such registration provides transparency in the review process, as all details about the process are defined in advance and accessible for peer reviewing. This ensues in a quality check and probable improvements of the review proposed prior to its performance. By revealing any differences among methods or outcomes reported in the published review relative to those planned in the registered protocol the rigor of the systematic review can be gauged. Vries et al. (2015) recommend documenting any revisions to the protocol within the materials and methods section of the systematic review, allowing readers to judge whether these may have introduced bias in the reviews. If this information is lacking the reader might assume that *e.g.* outcome parameters could have been intentionally omitted, resulting in publication bias. Finally, registration reduces unplanned duplication of systematic reviews (National Institute for Health Research NHS). Systematic reviews of animal studies are still not included in the Cochrane Library, but they can be registered on PROSPERO, which we did in the case described.

Team

The project started with three team members including one content expert (FR), one process expert (AP) and one project supervisor (CB). The inclusion of the content expert allows particularly to identify publication bias, but lead to additional challenges. In this case study, FR was part of the initial project team for vatinoxan hydrochloride, patent holder to the drug, and author of some of the studies which might be included in the final systematic review. To avoid risk of bias and guaranty impartiality, a review pattern was developed to prevent the author reviewing her own papers (compare section 3.7 point 22.). Soon, it became evident that further reviewers and experts for *e.g.*, literature research and statistics were needed to accomplish the work in a reasonable time and of high quality. O'Connor and Sargeant (2015) discussed the 'key-player' for conducting a systematic review in their review about research synthesis in veterinary science. In their paper they point out that it is advisable to integrate multiple experts (*e.g.*, librarians and statisticians) into the process according to needed skills and that reconsidering the membership and size of a systematic review team is often necessary (O'Connor and Sargeant 2015). Therefore, an expert on systematic reviews and meta-analysis

for animal studies was contacted [Assistant Professor PhD Carlijn Hooijmans (SYRCLE, RadboudUMC)] and additional experts at the Veterinary University of Vienna were invited to join the team (CV: meta-analysis and statistics, and DR: optimising search strategy).

Adding new members to the research team led to new challenges on how to train and integrate them into the existing team and how to transfer the knowledge. To face these challenges in the future it was agreed on a comprehensive description of all the tools that were created or adapted for the purpose of this review, as recommended by O'Connor et al. (2014b).

Amount of work

Although a systematic review requires substantially more resources and effort compared to a non-systematic review (BioMed Central Ltd, Springer Nature Systematic Reviews), it was decided to follow the above mentioned protocols and guidelines to avoid the scientific disadvantages. The structured process of planning the systematic review led to more focus than initially existed, showcased the need of teamwork to share expertise, and revealed the volume of data that should be collected in a single study to test for statistically and biologically significant differences (Huntley et al. 2017). While the process appears cumbersome, AP spent approximately 8 weeks on self-education about the available systematic review protocols and tools, along with terminology. Many protocol drafts were developed that were supported by regular in-person meetings. Extended discussions and careful revisions at this early stage are important to avoid changes to the protocol during the systematic review itself. A common understanding among members of the research team makes the process more repeatable to inform increasingly important decisions about clinical practice (O'Connor and Sargeant 2015).

Once a researcher or team is trained in the field of systematic reviews, developing protocols should become less time consuming. Based on the current experience, the importance of discussing and documenting critical terms (*e.g.* the outcome parameters) should not be underestimated. This is in line with the recommendation in qualitative risk analysis to limit linguistic uncertainty, which arises from the fact that words have different or imprecise meanings (Carey and Burgman 2008).

Conclusion

A variety of useful systematic review protocols, guidelines, and tools for health related clinical studies are available. There are, however, no specific ones for veterinary medicine. Modification of existing guidelines and tools are possible though. The SYRCLE protocol, designed for laboratory animal studies, is the closest template for the purpose of systematic reviews in clinical veterinary medicine. Independent of the protocol used, the consultation and inclusion of process experts is highly recommended, particularly if one is inexperienced in performing systematic reviews. Furthermore, awareness of time commitment is imperative to plan and perform a systematic review.

5. Abstract

The importance and awareness of evidence-based veterinary medicine (EVM) increased rapidly over the last decade. A cornerstone of EVM are systematic reviews. One of the most known and accepted authority for systematic reviews is the Cochrane Collaboration, which only covers clinical trials in human medicine. To the authors knowledge no specific protocols and guidelines exist for systematic reviews of clinical studies in veterinary medicine. The **objective** of this thesis was to provide an overview of selected systematic review protocols, guidelines and recommended tools. This was exemplified by the development and the preparation of a **systematic review protocol** including data extraction and evaluation template for a specific review question.

In recent years, **C**ollaborative **A**pproach to **M**eta-**A**nalysis and **R**eview of **A**nimal **D**ata from **E**xperimental **S**tudies (CAMRADES) and **SY**stematic **R**eview **C**enter for **L**aboratory animal **E**xperimentation (SYRCLE) developed guidelines and recommendations for systematic reviews on animal studies. Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA), **P**roblem/Patient/Population-**I**ntervention-**C**omparison-**O**utcomes-Approach (PICO), **G**rading of **R**ecommendations **A**ssessment, **D**evelopment and **E**valuation (GRADE) and **CON**solidated **S**tandards **O**f **R**eporting **T**rials (CONSORT) are some templates and tools which build the foundation of the Cochrane Handbook for systematic reviews and are reflected in the SYRCLE template as well. This thesis gives a brief overview of these guidelines/protocol templates and additional tools. The SYRCLE template could be used without modification in the present case study together with the aforementioned tools. The process of template selection, the refinement of the team and study question are described. Some challenges occurred during this process of which the definition of appropriate outcome parameters, judgement of quality of evidence and addressing the risk of bias were the main topics. Including further tools, adapting existing ones and adding new specialists to the team helped to overcome these challenges.

The SYRCLE protocol, designed for laboratory animal studies, was found to be the closest template for the purpose of systematic reviews in clinical veterinary medicine. Independent of the protocol used, the consultation and inclusion of process experts is highly recommended, particularly if one is inexperienced in performing systematic reviews. Furthermore, awareness of time commitment is imperative to plan and perform a systematic review.

6. Zusammenfassung

Im letzten Jahrzehnt ist die Wichtigkeit und Bekanntheit von evidenzbasierter Veterinärmedizin (EVM) deutlich angestiegen. Ein Eckpfeiler der EVM sind systematische Übersichtsarbeiten. Eine der bekanntesten und anerkanntesten Einrichtungen im Bereich der systematischen Übersichtsarbeiten ist die Cochrane Collaboration, die jedoch nur klinische Studien im Bereich der Humanmedizin abdeckt. Den Autoren sind aktuell keine spezifischen Protokolle oder Handbücher für systematische Übersichtsarbeiten im Bereich der Veterinärmedizin bekannt. Das Ziel dieser Arbeit war es, einen Überblick über ausgewählte Vorlagen für systematische Übersichtsarbeitsprotokolle, Handbücher und empfohlene Hilfsmittel, zu verschaffen. Dies sollte anhand eines konkreten Beispiels verdeutlich werden. In diesem wurde ein systematisches Übersichtsarbeitsprotokoll, inklusive einer Datenextraktions- und Evaluierungsvorlage, entwickelt und umgesetzt.

In den letzten Jahren haben „**C**ollaborative **A**pproach to **M**eta-**A**nalysis and **R**eview of **A**nimal **D**ata from **E**xperimental **S**tudies“ (CAMRADES) und „**SY**stematic **R**eview **C**enter for **L**aboratory animal **E**xperimentation“ (SYRCLE) Richtlinien und Empfehlungen für systematische Übersichtsarbeiten für Tierstudien entwickelt. „Preferred Reporting Items for Systematic Review and Meta-Analysis“ (PRISMA), „**P**roblem/Patient/Population-**I**ntervention-**C**omparison-**O**utcomes-Approach“ (PICO), „**G**rading of **R**ecommendations **A**ssessment, **D**evelopment and **E**valuation“ (GRADE) und „**CON**solidated **S**tandards **O**f **R**eporting **T**rials“ (CONSORT) sind einige Vorlagen und Hilfsmittel, welche die Grundlagen unter anderem für das Cochrane Handbuch für systematische Übersichtsarbeiten bilden und in der SYRCLE Vorlage berücksichtigt werden. Die vorliegende Arbeit gibt einen kurzen Überblick über diese Richtlinien/Protokollvorlagen und zusätzlichen Hilfsmittel. Die SYRCLE Vorlage konnte in der vorliegenden Studie, zusammen mit den zuvor genannten Hilfsmitteln, ohne Modifikation verwendet werden. Der Prozess der Vorlagenauswahl, das Zusammenstellen des Rezensions Teams und die Entwicklung der Fragestellung wurden beschrieben. Hierbei mussten einige Hürden überwunden werden. Die Auswahl geeigneter Beurteilungskriterien, die Bewertung der Qualität der Studienergebnisse und das Erkennen von Verzerrungsgefahren stellten hierbei die größten Herausforderungen dar. Diesen wurde durch die Integration weiterer Hilfsmittel, Anpassung bestehender Hilfsmittel und Hinzufügen von weiteren Experten zum Rezensions Team begegnet.

Das SYRCLE Protokoll, entwickelt für Labortierstudien, kommt einer Vorlage für systematische Übersichtsarbeiten im Bereich der Veterinärmedizin am nächsten. Unabhängig vom eingesetzten Protokoll wird unbedingt empfohlen Experten hinzuzuziehen, insbesondere, wenn man in diesem Bereich unerfahren ist. Des Weiteren ist es wichtig sich über den enormen Zeitaufwand für die Durchführung einer systematischen Übersichtsarbeit bewusst zu sein.

7. Abbreviations

AP	Andreas Palluch
BW	Body Weight
CAMRADES	Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies
CB	Dr.med.vet. Christina Braun Dipl.ACVAA
CH	Assistant Professor PhD Carlijn Hooijmans
CI	Confidence Interval
CO	Cardiac Output
CONSORT	CONsolidated Standards of Reporting Trials
CV	Dr.rer.nat. Priv.-Doz Claus Vogel
CVP	Central Venous Pressure
DAP	Diastolic Arterial Pressure
DR	Hofrätin Mag.med.vet. Doris Reinitzer
ECG	ElectroCardioGraphy
EP	Evidence Profile
EVM	Evidence-based Veterinary Medicine
FR	Dr.med.vet. Flavia Restitutti
GRADE	Grading of Recommendations Assessment Development and Evaluation
HR	Heart Rate
MAP	Mean Arterial Pressure
MD	Mean Difference
MVP	Mean Venous Pressure
NA	Not Available
OIS	Optimal Information Size
PICO	Population, Intervention, Comparators and Outcomes of interest
PIVA	Partial IntraVenous Anesthesia

PRISMA-P	Preferred Reporting Items for Systematic review and Meta-Analysis Protocols
PROSPERO	International Prospective Register of Systematic Reviews
RCT	Randomized Clinical Trial
RoB	Risk of Bias for animal intervention studies
SAP	Systolic Arterial Pressure
SMD	Standardized Mean Difference
SOF	Summary Of Findings table
SVR	Systemic Vascular Resistance
SVRI	Systemic Vascular Resistance Index
SYRCLE	SYstematic Review Center for Laboratory animal Experimentation
TIVA	Total IntraVenous Anesthesia
TS	Tamara Schütz

8. References

Ackley BJ. 2008. Evidence-based nursing care guidelines. Medical-surgical interventions. St. Louis, Mo.: Mosby/Elsevier, 1 online resource.

BioMed Central Ltd, Springer Nature Systematic Reviews. https://systematicreviewsjournal.biomedcentral.com/ (accessed Sep 14, 2018).

Carey JM, Burgman MA. 2008. Linguistic uncertainty in qualitative risk analysis and how to minimize it. Annals of the New York Academy of Sciences, 1128: 13–17. DOI 10.1196/annals.1399.003.

Clarke KW, Hall LW. 1969. "Xylazine"--a new sedative for horses and cattle. The Veterinary record, 85 (19): 512–517.

Cochrane. https://www.cochranelibrary.com/ (accessed Apr 17, 2019).

Enouri SS, Kerr CL, McDonell WN, O'Sullivan ML, Neto FJT. 2008. Effects of a peripheral alpha2 adrenergic-receptor antagonist on the hemodynamic changes induced by medetomidine administration in conscious dogs. American journal of veterinary research, 69 (6): 728–736. DOI 10.2460/ajvr.69.6.728.

Gopalakrishnan S, Ganeshkumar P. 2013. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. Journal of family medicine and primary care, 2 (1): 9–14. DOI 10.4103/2249-4863.109934.

Guyatt GH, Ebrahim S, Alonso-Coello P, Johnston BC, Mathioudakis AG, Briel M, Mustafa RA, Sun X, Walter SD, Heels-Ansdell D, Neumann I, Kahale LA, Iorio A, Meerpohl J, Schünemann HJ, Akl EA. 2017. GRADE guidelines 17: assessing the risk of bias associated with missing participant outcome data in a body of evidence. Journal of clinical epidemiology, 87: 14–22. DOI 10.1016/j.jclinepi.2017.05.005.

Guyatt GH, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. 2011b. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of clinical epidemiology, 64 (4): 383–394. DOI 10.1016/j.jclinepi.2010.04.026.

Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schünemann HJ. 2011g. GRADE guidelines: 2. Framing the question and deciding

on important outcomes. Journal of clinical epidemiology, 64 (4): 395–400. DOI 10.1016/j.jclinepi.2010.09.012.

Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G, Jaeschke R, Williams JW, Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schünemann HJ. 2011e. GRADE guidelines 6. Rating the quality of evidence--imprecision. Journal of clinical epidemiology, 64 (12): 1283–1293. DOI 10.1016/j.jclinepi.2011.01.012.

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M, Schünemann HJ. 2011d. GRADE guidelines: 8. Rating the quality of evidence--indirectness. Journal of clinical epidemiology, 64 (12): 1303–1310. DOI 10.1016/j.jclinepi.2011.04.014.

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA, Norris S, Vist G, Dahm P, Shukla VK, Higgins J, Falck-Ytter Y, Schünemann HJ. 2011c. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. Journal of clinical epidemiology, 64 (12): 1294–1302. DOI 10.1016/j.jclinepi.2011.03.017.

Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, Williams JW, Meerpohl J, Norris SL, Akl EA, Schünemann HJ. 2011a. GRADE guidelines: 5. Rating the quality of evidence--publication bias. Journal of clinical epidemiology, 64 (12): 1277–1282. DOI 10.1016/j.jclinepi.2011.01.011.

Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, Atkins D, Kunz R, Brozek J, Montori V, Jaeschke R, Rind D, Dahm P, Meerpohl J, Vist G, Berliner E, Norris S, Falck-Ytter Y, Murad MH, Schünemann HJ. 2011f. GRADE guidelines: 9. Rating up the quality of evidence. Journal of clinical epidemiology, 64 (12): 1311–1316. DOI 10.1016/j.jclinepi.2011.06.004.

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW, Atkins D, Meerpohl J, Schünemann HJ. 2011h. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of

bias). Journal of clinical epidemiology, 64 (4): 407–415. DOI 10.1016/j.jclinepi.2010.07.017.

Haidich AB. 2010. Meta-analysis in medical research. Hippokratia, 14 (Suppl 1): 29–37.

Higgins JPT. 2008. Cochrane handbook for systematic reviews of interventions. First ed. Chichester: Wiley-Blackwell, 649.

Honkavaara JM, Raekallio MR, Kuusela EK, Hyvärinen EA, Vainio OM. 2008. The effects of L-659,066, a peripheral alpha2-adrenoceptor antagonist, on dexmedetomidine-induced sedation and bradycardia in dogs. Veterinary anaesthesia and analgesia, 35 (5): 409–413. DOI 10.1111/j.1467-2995.2008.00403.x.

Hooijmans CR, Rovers MM, Vries RBM de, Leenaars M, Ritskes-Hoitinga M, Langendam MW. 2014. SYRCLE's risk of bias tool for animal studies. BMC medical research methodology, 14: 43. DOI 10.1186/1471-2288-14-43.

Huntley SJ, Dean RS, Brennan ML. 2017. The Awareness of the International Veterinary Profession of Evidence-Based Veterinary Medicine and Preferred Methods of Training. Veterinary sciences, 4 (1). DOI 10.3390/vetsci4010015.

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. 1996. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Controlled clinical trials, 17 (1): 1–12.

Jüni P, Altman DG, Egger M. 2001. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ (Clinical research ed.), 323 (7303): 42–46. DOI 10.1136/bmj.323.7303.42.

Leenaars M, Hooijmans CR, van Veggel N, ter Riet G, Leeflang M, Hooft L, van der Wilt GJ, Tillema A, Ritskes-Hoitinga M. 2012. A step-by-step guide to systematically identify all relevant animal studies. Laboratory animals, 46 (1): 24–31. DOI 10.1258/la.2011.011087.

Light RJ, Pillemer DB. 1984. Summing up. The science of reviewing research. Cambridge, Mass: Harvard University Press, 191.

Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. 2010. CONSORT 2010 explanation and elaboration: updated guidelines

for reporting parallel group randomised trials. BMJ (Clinical research ed.), 340: c869. DOI 10.1136/bmj.c869.

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews, 4: 1. DOI 10.1186/2046-4053-4-1.

Muir WW, Ueyama Y, Noel-Morgan J, Kilborne A, Page J. 2017. A Systematic Review of the Quality of IV Fluid Therapy in Veterinary Medicine. Frontiers in veterinary science, 4: 127. DOI 10.3389/fvets.2017.00127.

Murdoch University. https://libguides.murdoch.edu.au/ld.php?content_id=19818940 (accessed Mar 31, 2019).

Murrell JC, Hellebrekers LJ. 2005. Medetomidine and dexmedetomidine: a review of cardiovascular effects and antinociceptive properties in the dog. Veterinary anaesthesia and analgesia, 32 (3): 117–127. DOI 10.1111/j.1467-2995.2005.00233.x.

National Institute for Health Research NHS. https://www.crd.york.ac.uk/prospero/#aboutpage (accessed Sep 13, 2018).

O'Connor A, Sargeant J. 2015. Research synthesis in veterinary science: Narrative reviews, systematic reviews and meta-analysis. Veterinary journal, 206 (3): 261–267. DOI 10.1016/j.tvjl.2015.08.025.

O'Connor AM, Anderson KM, Goodell CK, Sargeant JM. 2014b. Conducting systematic reviews of intervention questions I: Writing the review protocol, formulating the question and searching the literature. Zoonoses and public health, 61 Suppl 1: 28–38. DOI 10.1111/zph.12125.

O'Connor AM, Sargeant JM, Wang C. 2014a. Conducting systematic reviews of intervention questions III: Synthesizing data from intervention studies using meta-analysis. Zoonoses and public health, 61 Suppl 1: 52–63. DOI 10.1111/zph.12123.

Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, Britten N, Roen K, Duffy S. 2006. Guidance on the conduct of narrative synthesis in systematic reviews. Results of an ESRC funded research project. Unpublished report. Lancaster.

Raekallio MR, Honkavaara JM, Vainio OM. 2010. The effects of L-659,066, a peripheral α2-adrenoceptor antagonist, and verapamil on the cardiovascular influences of dexmedetomidine in conscious sheep. Journal of veterinary pharmacology and therapeutics, 33 (5): 434–438. DOI 10.1111/j.1365-2885.2009.01156.x.

Richardson WS, Wilson MC, Nishikawa J, Hayward RS. 1995. The well-built clinical question: a key to evidence-based decisions. ACP journal club, 123 (3): A12-3.

Saaiq M, Ashraf B. 2017. Modifying "Pico" Question into "Picos" Model for More Robust and Reproducible Presentation of the Methodology Employed in A Scientific Study. World journal of plastic surgery, 6 (3): 390–392.

Sampson M. 2014. Welcoming systematic reviews to the Journal of the Medical Library Association. Journal of the Medical Library Association : JMLA, 102 (3): 143–145. DOI 10.3163/1536-5050.102.3.001.

Sargeant JM, O'Connor AM. 2014. Conducting systematic reviews of intervention questions II: Relevance screening, data extraction, assessing risk of bias, presenting the results and interpreting the findings. Zoonoses and public health, 61 Suppl 1: 39–51. DOI 10.1111/zph.12124.

Schulz KF, Altman DG, Moher D. 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ (Clinical research ed.), 340: c332. DOI 10.1136/bmj.c332.

Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ (Clinical research ed.), 350: g7647. DOI 10.1136/bmj.g7647.

Sinclair MD. 2003. A review of the physiological effects of alpha2-agonists related to the clinical use of medetomidine in small animal practice. The Canadian veterinary journal, 44 (11): 885–897.

The University of Toledo. http://libguides.utoledo.edu/sysrev/collaboration (accessed Mar 27, 2020). (last updated Mar 27, 2020).

University of Nottingham, Center of Evidence-based veterinary medicine. https://www.nottingham.ac.uk/cevm/about-the-cevm/evidence-based-veterinary-medicine-%28evm%29.aspx (accessed Sep 7, 2018).

US National Library of Medicine National Institutes of Health. https://www.ncbi.nlm.nih.gov/pubmed/?term=veterinary+medicine+systematic+review (accessed Apr 12, 2019).

Vainionpää MH, Raekallio MR, Pakkanen SAE, Ranta-Panula V, Rinne VM, Scheinin M, Vainio OM. 2013. Plasma drug concentrations and clinical effects of a peripheral alpha-2-adrenoceptor antagonist, MK-467, in horses sedated with detomidine. Veterinary anaesthesia and analgesia, 40 (3): 257–264. DOI 10.1111/vaa.12012.

van Loveren C, Aartman IHA. 2007. De PICO-vraag. Nederlands tijdschrift voor tandheelkunde, 114 (4): 172–178.

Vesterinen HM, Sena ES, Egan KJ, Hirst TC, Churolov L, Currie GL, Antonic A, Howells DW, Macleod MR. 2014. Meta-analysis of data from animal studies: a practical guide. Journal of neuroscience methods, 221: 92–102. DOI 10.1016/j.jneumeth.2013.09.010.

Vries RBM de, Hooijmans CR, Langendam MW, van Luijk J, Leenaars M, Ritskes-Hoitinga M, Wever KE. 2015. A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies. Evidence-based Preclinical Medicine, 2 (1): e00007. DOI 10.1002/ebm2.7.

Zeiler GE. 2015. A review of clinical approaches to antagonism of alpha 2-adrenoreceptor agonists in the horse. Equine Veterinary Education, 27 (1): 48–54. DOI 10.1111/eve.12249.

9. List of figures and tables

10. Appendix 1: GRADE level of evidence and data extraction template/example

Study title:
Author:
Declaration of interest:
Funding source:
Reason for exclusion:

Level of overall trustworthiness:
Journal:
received:
accepted:
published:

Evidence Profile

Study design: Randomized trial

Outcome/ Observed	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication bias	Large effect	Dose response	All plausible residual confounding	Quality of a body of evidence		Justification for Grading	Importance
attenuation of bradycardia					>							
no	0	0	0	0	0	0	0	0	N/A	no evidence level		8 - critical
maintenance of cardiac output (CO)					>							
no	0	0	0	0	0	0	0	0	N/A	no evidence level		8 - critical
prevention of SVR increase					>							
no	0	0	0	0	0	0	0	0	N/A	no evidence level		7 - critical
maintenance of mean arterial pressure (MAP)					>							
no	0	0	0	0	0	0	0	0	N/A	no evidence level		8 - critical

primary outcome parameter | secondary outcome parameter

Adverse effects:
Other indication:
Pharmacokinetic:
Power calculation:

Data extraction:
Aim/Hypothesis:

Materials and Method:

Study design:	blinded?	randomized?	cross-over?	well describe	comment:
animals:	species:	number total:		age median:	weight median:
		male:		age range:	weight range:
	breed:	female:		age mean:	weight mean:
		mk:		age sd:	weight sd:
which?		fs:			
condition:		anasthesia:			
comment:					

Drugs:	alpha2agonist	dose (mg/kg):	dose (µg/kg):	route:	placebo (NaC comment:
1					

2						
3						

vatinoxan:	dose (mg/kg):	dose (µg/kg):	route:	Timepoint:	number of pa	recommende	comment:
dose1:							
dose2:							
dose3:							
dose4:							
dose5:							

Protocols comment:

Further Outcomes:

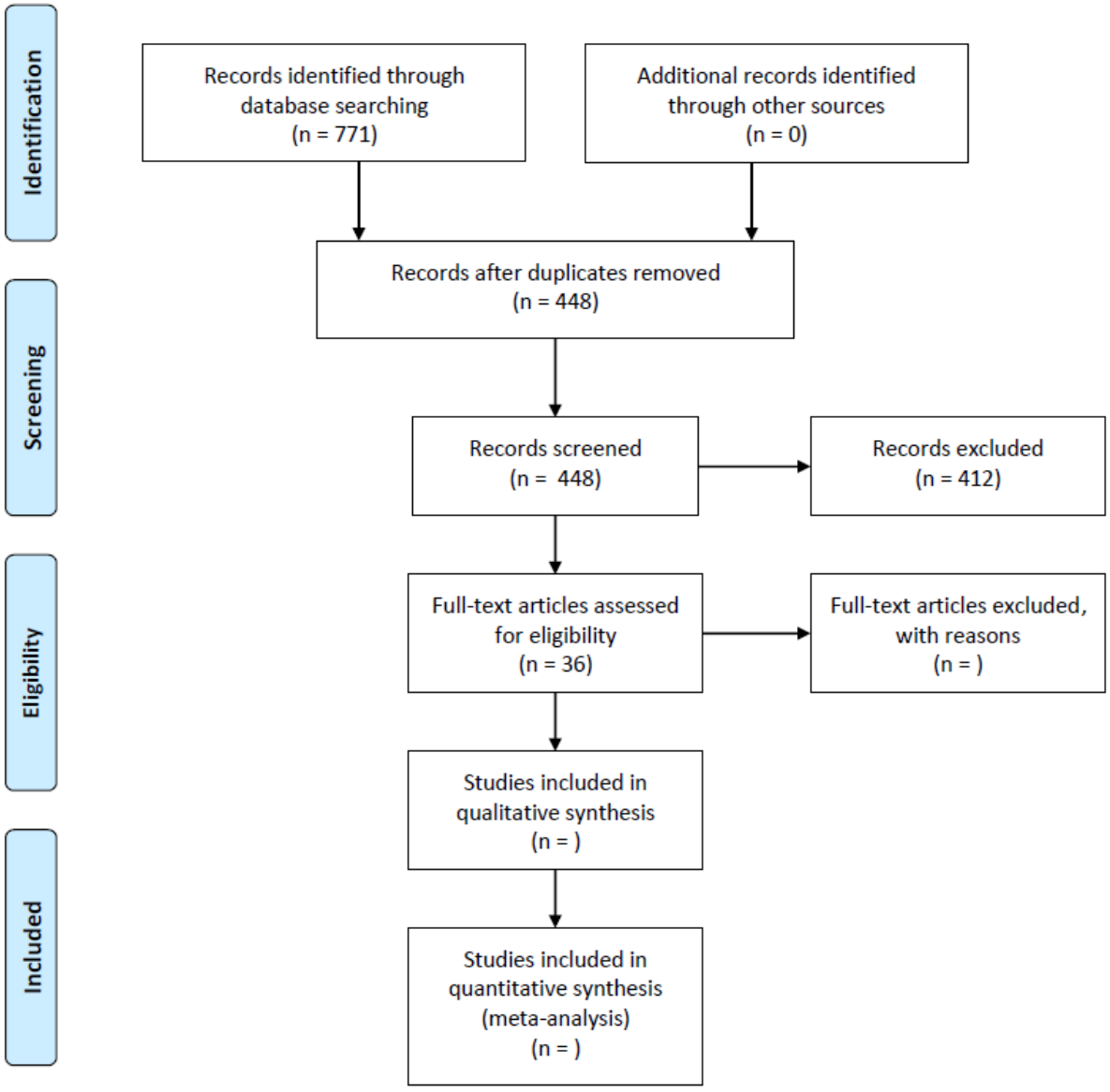
	Measured?	Effect?	Direction:	Magnitude:	Method	Comment:
Sedation:						
Analgesia:						
Blood gases:						
Respiratory rate:						
Recovery quality:						
GIT:						
Glucose/Insuline:						

Results:

Discussion:

Conclusion:

11. Appendix 2: Template for PRISMA Flow Diagram



(Shamseer et al. 2015)