

EXTERNAL SCIENTIFIC REPORT

Preparatory work for the development of a scientific opinion on the main welfare risks related to the farming of sheep for wool, meat and milk production¹

A. O'Connor¹, D. Wolfe², J. Sargeant², J. Glanville³ and H. Wood³

¹Iowa State University, Ames Iowa, USA, ²University of Guelph, Guelph, CA, ³York Health Economics Consortium, University of York, UK

ABSTRACT

This report contains the results of a scoping review of sheep welfare studies and a systematic review of the effect of extensive/outdoor/migratory management on lameness compared to intensive/indoor management systems in sheep raised for the production of meat, milk, or wool in Europe. The scoping review allowed identifying and mapping 679 citations relevant to sheep welfare. Those citations were mapped according to the study population, 8 main welfare determinants (management, environment, genetics, nutrition/feeding/watering, behaviour, health, housing, handler traits/human-animal bond) and outcomes. Such mapping supported the WG in identifying gaps of knowledge and data that further led to seeking for experts' knowledge, as well as to identify areas where a systematic literature process could be performed. The systematic review that followed the scoping review provided evidence that the management system is not associated with the prevalence or risk of lameness. However, higher stocking densities were associated with prevalence or risk lameness. The body of work may appear to be quite small, however, given the difficulties faced by researchers investigating this topic, identifying studies that looked at these factors in a limited region of the world is a reasonable body of work. This is a difficult topic to study as the exposures are variable and the outcomes difficult to measure in production systems as they can occur year round and have numerous causes.

© European Food Safety Authority, 2014

KEY WORDS

welfare, sheep, production, lameness, interdigital dermatitis, scoping review, systematic review

¹ Question No EFSA-Q-2013-00580.

Any enquiries related to this output should be addressed to ALPHA@efsa.europa.eu

Suggested citation: O'Connor et al, 2014. Preparatory work for the development of a scientific opinion on the main welfare risks related to the farming of sheep for wool, meat and milk production. EFSA supporting publication 2014-EN-678, 93 pp.

Available online: www.efsa.europa.eu/publications

DISCLAIMER

The present document has been produced and adopted by the bodies identified above as authors. This task has been carried out exclusively by the authors in the context of a contract between the European Food Safety Authority and the authors, awarded following a tender procedure. The present document is published complying with the transparency principle to which the Authority is subject. It may not be considered as an output adopted by the Authority. The European food Safety Authority reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors

TABLE OF CONTENTS

Abstract	1
Table of contents	3
Background as provided by efsa	6
Terms of reference as provided by efsa.....	6
Introduction and Objectives	7
A scoping review of sheep welfare studies	9
Materials and Methods	9
1. Identification of literature for relevance screening and data mapping	9
1.1. Identification of an initial literature set.....	10
1.2. Grouping of citations into clusters	11
1.3. Identification of clusters of interest	11
1.4. Identification of the final body of literature for the scoping exercise.....	11
2. Relevance screening and data mapping	11
2.1. Development and refinement of a set of questions for relevance screening and characterization	11
2.1.1. Data mapped in relevant citations.....	12
2.1.1.1. Welfare determinants	12
2.1.1.2. Welfare outcomes	12
2.1.1.3. Characteristics of study or study population.....	13
2.1.1.4. Language of full publication and country of study	14
2.2. Second-level data characterization.....	17
2.3. Relevance screening and data mapping methods.....	17
3. Identification of relevant questions for systematic review	18
3.1. Teleconference with EFSA Working Group to discuss D2 deliverable and agree upon a relevant question(s) for systematic review.....	18
3.2. Secondary non-comprehensive scoping exercise specific to sheep lameness.....	18
3.2.1. Search methods for secondary sheep lameness scoping exercise.....	18
3.2.2. Relevance screening and data mapping for secondary sheep lameness scoping exercise	18
3.3. Secondary non-comprehensive scoping exercise specific to sheep body condition score....	20
3.3.1. Search methods for secondary sheep body condition score scoping exercise	20
3.3.2. Relevance screening and data mapping for secondary sheep lameness scoping exercise	20
3.4. Follow-up teleconference to discuss findings of the secondary scoping exercises and to agree upon a relevant question(s) for systematic review	20
Results	21
4. Identification of literature for relevance screening and data mapping	21
5. Relevance screening and data mapping	22
5.1. Data mapping.....	23
5.1.1. Welfare determinants.....	23
5.1.2. Welfare outcomes	24
5.1.3. Characteristics of the study or study population.....	25
5.1.4. Language of full publication and country of study	25
6. Identification of relevant questions for systematic review	26
6.1. Secondary non-comprehensive scoping exercise specific to sheep lameness.....	27
6.2. Secondary non-comprehensive scoping exercise for sheep body condition score.....	28
6.3. Follow-up teleconference to discuss findings of the secondary scoping exercises and to agree upon a relevant question(s) for systematic review	28
Conclusions	28

A systematic review of the effect of extensive/outdoor/migratory management on lameness compared to intensive/indoor management systems in sheep raised for the production of meat, milk, or wool in europe	30
Introduction and Objectives	30
Materials and methods.....	30
6.4. PECO(S) Question	30
6.5. Protocol and registration.....	30
6.6. Eligibility criteria.....	30
6.6.1. Relevant populations	30
6.6.2. Relevant exposure or comparisons.....	31
6.6.3. Types of outcome measures	31
6.6.4. Relevant study designs	31
6.7. Information Sources.....	31
6.8. Search strategy	32
6.8.1. Search terms	32
6.8.2. Search considerations	32
6.8.3. Search results.....	32
6.9. Study Selection	32
6.9.1. Screening for eligibility for the review.....	33
6.10. Data collection process	33
6.11. Data Items	33
6.11.1. General study characteristics	33
6.11.2. Outcome-level and exposure information	33
6.11.3. Exposure information	33
6.12. Assessment of risk bias in included studies	34
6.13. Summary measures	34
6.14. Data synthesis	34
6.15. Data synthesis	34
6.16. Additional Analyses.....	34
Results	34
6.17. Search results	34
6.18. Study selection.....	35
6.18.1. Screening for eligibility for the review.....	35
6.19. Study characteristics	35
6.19.1. General study characteristics (potential sources of clinical heterogeneity).....	35
6.19.2. Exposure level information	35
6.19.3. Outcome-level information.....	36
6.19.3.1. Management system - extensive/outdoor/migratory management (E) or intensive/indoor management.....	36
6.19.3.2. Stocking density.....	36
6.20. Assessment of risk of bias within studies	36
6.21. Results of risk of bias within studies	37
6.21.1. Risk of bias across studies	38
6.22. Summary of evidence	38
6.23. Limitations.....	38
Conclusions	38
References	55
Appendices	56
Appendix for the scoping review	56
Appendix A. Refinement of literature search for scoping review.....	56

Refinement of the search terms to be used for identification of the literature set for the scoping review—developing a full search strategy	56
Appendix B. The abstract screening and abstract characterization questions for scoping review...	59
Appendix C. Table describing the number of studies in each welfare determinate and the frequency of outcomes reported in that determinate for scoping review.	62
Appendix for the systematic review	66
Appendix D. Search strategies for systematic review	66
Appendix E. Study Selection questions for systematic review	74
Appendix F. Manuscript level questions for data extraction and clarification notes for systematic review	75
Appendix G. Outcome and exposure level questions and notes for data extraction for systematic review	78
Appendix H. Risk of bias form and interpretation notes for systematic review	82
Appendix I. Proposed protocol for review of the literature evaluating the effect of management system on lameness in sheep raised for the production of meat, milk, or wool in Europe	85

BACKGROUND AS PROVIDED BY EFSA

The Panel on Animal Health and Welfare (AHAW) provides scientific advice on all aspects of animal disease and animal welfare. Its work chiefly concerns food producing animals, including fish.

The Panel carries out risk assessments in order to produce scientific opinions and advice for risk managers. Its risk assessment approach is based on reviewing scientific information and data in order to evaluate the risks as a consequence of a given hazard. This helps to provide a science-based foundation for European policies and legislation, and supports risk managers in taking balanced and timely decisions.

While no specific EU rules on farming of sheep exist, the EU Strategy for the protection and welfare of animals 2012-2015 foresees a revised animal welfare framework, introducing the use of animal-based welfare indicators to simplify the legal framework and to enhance the applicability of general principles to all farm animals.

Meanwhile international organisations, global stakeholders and Third Countries Governments are moving towards more sustainable livestock production policies and farming practices, developing guidelines and codes of practices addressing the welfare of sheep. This includes the recent joint initiative of the Commission (DG SANCO) and the International Wool and Textile Organisation (IWTO) to support the elaboration of a guideline for best practices for welfare of wool producing animals.

The work being carried out by the AHAW Panel on the assessment of risks for animal welfare and on the use of animal-based measures is supporting the Commission in the preparation of this new legal framework.

The EFSA AHAW Unit intends to launch a reopening competition under this framework contract in order to commission a systematic review on the welfare of sheep. The review questions stem from the terms of reference of the mandate, which is available on the EFSA Register of Questions (EFSA/Q/2013/00850)

TERMS OF REFERENCE AS PROVIDED BY EFSA

This contract/grant was awarded by EFSA to:

Contractor/Beneficiary: Annette O'Connor at the Iowa State University, Julie Glanville at the York Health Economics Consortium, University of York and Jan Sargeant at the University of Guelph.

Contract/grant title: "Preparatory work for the development of a scientific opinion on the main welfare risks related to the farming of sheep for wool, meat and milk production"

Contract/grant number: RC/EFSA/AHAW/2013/03

INTRODUCTION AND OBJECTIVES

This specific document is the final report summarising the results of the a scoping review and systematic review conducted for the EFSA project “Preparatory work for the development of a scientific opinion on the main welfare risks related to the farming of sheep for wool, meat and milk production, Call Reference: RC/EFSA/AHAW/2013/03.”

The document is presented with the processes and results of the scoping review 1st followed by the processes and result of the systematic review.

The 1st part of this report summarises the results of the start-up contract for the EFSA project “Preparatory work for the development of a scientific opinion on the main welfare risks related to the farming of sheep for wool, meat and milk production, Call Reference: RC/EFSA/AHAW/2013/03.” The overall purpose of this project was to characterize the welfare determinates and outcomes that have been studied by welfare researchers with respect to sheep involved in the production of wool, meat, and milk. Based on knowledge of the topics that have been studied (reported in the D2 deliverable for this project), a protocol for systematic review of a suitable topic(s) will be included as part of this final deliverable for this start-up contract..

The tasks to be completed by the end of this specific start-up contract as listed in the technical specifications document were

Task 1:

Sub-task 1: Identification and formulation of the relevant review question(s) necessary to address the contract terms of reference. The task will be developed based on initial discussion with EFSA and accurate scoping exercises conducted by the contractor as part of the start-up specific contract.

Sub-task 2: Mapping of the evidence found through the scoping exercises by study design and risk factors / welfare consequences / animal-based measures identified.

Sub-task 3: Preparation of the list of papers resulting from the accurate scoping exercises (Endnote library).

Task 2:

Preparation of a literature review protocol for the questions identified in Task 1 and agreed with EFSA. The protocol should describe in detail the search equations and electronic databases to be searched as well as any other sources of information to be included, the reference management system to be used, the relevance criteria, the eligibility criteria, roles and tasks of reviewers, methodology for conflict resolution between reviewers and data extraction and analysis methodology.

The task to be completed for the follow-up specific contract, as specified in the technical specifications was

Task 3:

Production of systematic review(s) for the agreed review question(s), reported in an appropriate format.

Specifically, this final report presents the approach to and results of activities designed to achieve Tasks 1 and 2. To achieve those tasks the following objectives were identified:

1. Identification of an initial literature set for the scoping review
2. Grouping of citations into clusters, based on terms used within the initial literature set
3. Identification of clusters of interest and apparently relevant terms used therein
4. Identification of the final body of literature for the scoping exercise through refinement of search terms
5. Development and refinement of a set of questions to screen and to map the relevant data found in the literature set
6. Screening of identified literature for relevance to on-farm sheep welfare
7. Mapping of data within relevant studies
8. Discussion of findings of data mapping exercise and proposed review questions with the EFSA sheep welfare working group to agree upon the relevant question(s) for which a protocol(s) would be subsequently drafted
9. Development of systematic review protocol(s) for the agreed upon questions

A SCOPING REVIEW OF SHEEP WELFARE STUDIES

MATERIALS AND METHODS

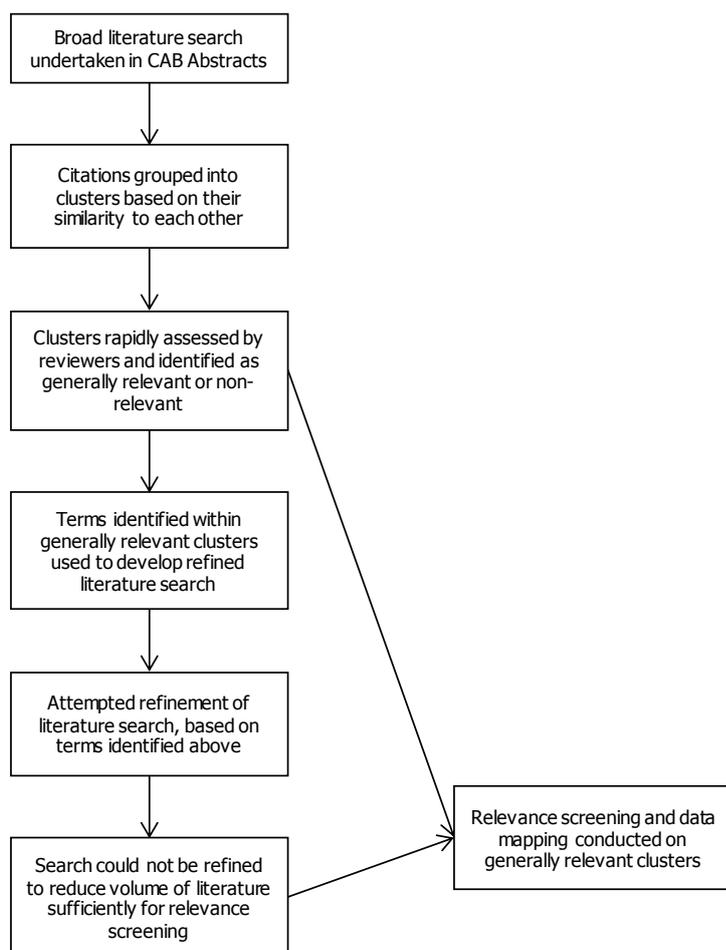
1. Identification of literature for relevance screening and data mapping

The literature relevant to on-farm sheep welfare is very large and covers a number of disciplines and study designs. Moreover, it is difficult to construct a search with sufficient sensitivity to capture the key element of welfare because the free-text and index terms that could be used to describe welfare issues are extremely diverse, are not consistently applied across relevant studies, and are non-specific to a welfare context. A search strategy based on such a wide range of terms would return extremely large numbers of results. For example, stress and behavioural responses are frequently used as indicators of sheep welfare; however, “stress” and “behavior(al)/behaviour(al)” are very non-specific and retrieve large volumes of results on non-relevant topics. Given the time constraints imposed by the project deadlines, it was not possible to screen a large volume of studies derived from a systematic search with an appropriate balance between sensitivity (identifying all relevant records) and specificity (limiting the volume of irrelevant records retrieved).

A multi-step process was used to identify a body literature that could be screened for relevance in the time available (Figure 1). Initially, a very broad literature search was conducted, and the citations identified by this search were categorized into clusters and visualised using text-mining software (Biowisdom, Version 6.0.3). Text-mining software is able to interrogate a large, representative corpus of literature concerned with sheep welfare and identify “clusters” of records which illustrate the most prevalent topics and terminology within the corpus. The rationale of this initial approach was to provide a sense of the scale and scope of the literature, and identify topics and search terms to inform the final search strategy, using a sample over 10 times larger than we can usually assess using traditional methods.

The clusters identified by the text mining software were then manually and rapidly evaluated for general relevance to on-farm sheep welfare by the reviewers, with the intent that terms identified within the relevant clusters by the text-mining software could be used to refine the original literature search, resulting in a volume of literature of a size suitable for relevance screening in the available time.

Figure 1: Process used to identify literature for relevance screening and data mapping



1.1. Identification of an initial literature set

The initial broad search undertaken in CAB Abstracts (via Web of Knowledge) is presented in Figure 2. CAB Abstracts was selected as information retrieval research suggests that this resource has the strongest coverage of journal titles in the field of veterinary science, agriculture, and animal health.^{2,3} The search was not limited by language, study design, or date in order to retrieve the most representative set of records possible.

Figure 2: Search strategy designed for CAB Abstracts (via Web of Knowledge, Thomson Reuters). 1910 to latest update. Run 10 January 2014.

TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton") AND ("welfare" OR "behavior" OR "behaviour" OR "behaviors" OR "behaviours" OR "behavioral" OR "behavioural" OR "stress" OR "stressor*") 13,194 results

² Grindlay DJC, Brennan ML, Dean RS, 2012. Searching the veterinary literature: a comparison of the coverage of veterinary journals by nine bibliographic databases. *Journal of Veterinary Medical Education*; 39, 4, 404-412.

³ Kawasaki J, 2004. Agriculture journal literature indexed in life sciences databases. *Issues in Science and Technology Librarianship*, 40.

1.2. Grouping of citations into clusters

Citations identified by the initial search strategy were imported into OmniViz data analysis and visualisation software (Biowisdom, Version 6.0.3). The records were clustered by the software using the K-Means method, whereby the abstracts were partitioned into a predefined number of clusters. This is recommended for large data sets in excess of 3,000 records. Exploratory work suggested that clustering using the content of the abstract database field provided the most useful results, rather than clustering based on title, keywords, or a combination of title, abstract, and keywords.

Each cluster was labelled with some of the most important terms that influenced the clustering: the terms the software algorithm identified as best at grouping the cluster of citations. The citations making up each cluster, and their labels were downloaded and passed to reviewers in Iowa (ISU) and Guelph for assessment.

1.3. Identification of clusters of interest

The review team rapidly assessed the abstracts of the citations within the clusters to identify clusters that appeared to have a large number of citations relevant to on-farm sheep welfare. The approach to rapid assessment was to scan the first and last 10 abstracts within each cluster for potential relevance. This was not a detailed analysis of the abstracts. Clusters that were considered to have a relatively high prevalence of potentially relevant citations were then identified as clusters of interest.

The results of the cluster analysis were presented back to the information specialists at York University for refinement of the literature search terms.

1.4. Identification of the final body of literature for the scoping exercise

The terms associated with the clusters of interest identified by the reviewers were assessed for their utility in refinement of the original literature search. Because further refinement of the search terms could not sufficiently limit the volume of literature identified for relevance screening (see Appendix A), the review team elected to conduct relevance screening on the citations contained within the clusters of interest identified by the reviewers.

2. Relevance screening and data mapping

2.1. Development and refinement of a set of questions for relevance screening and characterization

Concurrent to the identification of the body of literature for the scoping exercise, a set of questions was developed by the team at ISU and Guelph to screen citations for relevance and to broadly characterize the information within the relevant citations.

Relevant citations were those that described a factor that potentially influenced the on-farm welfare of sheep used in wool, meat, or milk production. After identifying relevant citations, characterization based only on the abstract of the citation was conducted. For this project, characterisation or mapping refers to process of extracting pertinent information from the abstracts of relevant citations to create a virtual map of topics studied in the screened literature set. This approach might also be referred to as scoping (i.e., to understand the scope of topics studied in the screened literature set). To characterize citations it is necessary to develop a series of questions that serve to extract the requested data. The characterization of the citation was only conducted for these relevant studies. Potentially relevant reviews/books/reports and conference proceedings were also identified, as well as potentially relevant studies for which limited detail in the abstract prevented proper screening. The topics covered by these citations were not mapped.

Question development for the characterization was an iterative process, with an initial set of questions refined after detailed scrutiny of the EFSA document “*DRAFT table of main welfare consequences and factors Rev 8*” (EFSA Working Group on the welfare of sheep) and after question testing by the ISU/Guelph team. A draft set of questions was provided to the EFSA Working Group on the welfare of sheep for testing at their meeting on 14 January 2014, and further revisions were made based upon comments from the Working Group. Following testing of this revised set of questions by the ISU/Guelph team, further revisions were completed, with the changes being presented to the EFSA Working Group contacts for comment on 17 January 2014. Question revision continued iteratively, based upon comments made by the EFSA Working Group contacts and upon testing of the questions, until a final set of questions was developed.

2.1.1. Data mapped in relevant citations

2.1.1.1. Welfare determinants

Welfare determinants were defined as those factors studied or reported as being potentially associated with a change in sheep welfare on farm. The “principles” identified by the EFSA Working Group on the welfare of sheep in the EFSA document “*DRAFT table of main welfare consequences and factors Rev 8*” (EFSA Working Group on the welfare of sheep)—i.e., feeding, housing/environment, health, and behaviour—were modified to create an initial list of welfare determinants. The final set of welfare determinants used for characterization are presented in Table 1. During the data mapping exercise, multiple options could be chosen during the characterization process i.e. a citation may have addressed multiple welfare determinants.

The genetics determinant was defined as including studies that looked at genetic loci, heritability indexes and breed. The welfare determinant “nutrition/feeding/watering” was defined as encompassing all management related to oral intake by sheep (i.e., food type, food/water volume intake, trough types, water source types, supplementation, etc.). The housing category was used for all structural components of the indoor environment that were not related to food or water or ventilation (e.g., flooring, pen construction, bedding). The environment category included all non-structural aspects of indoor environments that were not related to food or water (e.g., ventilation, temperature, sound, lighting, stocking density) and all aspects of outdoor environments not related to food or water (e.g., terrain, temperature, shade, stocking density). Management-related welfare determinants were management practices not related to those described under the nutrition/feeding/watering, housing, or environment categories, and included comparisons of methods for castration, weaning, shearing; evaluations of the effects of isolation and group confinement, etc. Health-related welfare determinants included specific diseases that could potentially influence sheep welfare (e.g., parasitism, dystocia). Similarly, behaviour-related welfare determinants included specific behaviours that could influence sheep welfare (e.g., mis-mothering, aggression/bullying). The handler traits/human-animal bond category encompassed the influence of handler competence and gentling techniques on sheep welfare.

2.1.1.2. Welfare outcomes

A final list of 33 outcomes was created, initially based upon the “welfare criteria” and “negative welfare consequences” presented in the EFSA document “*DRAFT table of main welfare consequences and factors Rev 8*” (EFSA Working Group on the welfare of sheep). These outcomes were modified and an inclusive list of the measurable welfare indicators that could be reported in relevant studies was developed (Table 1). Multiple options could be chosen during the mapping process. Some categories were included based on recommendation of the EFSA staff of working group, although they seemed to overlap with other outcomes or to be poorly defined. As a consequence, it was decided for the outcomes “distress,” “pain,” and “fear,” these would be used only when the citation included these explicit terms. Furthermore, for the terms “stress” and “thermal stress,” it was decided to report these

terms when the authors did not provide other outcomes (i.e., no measureable indicator of stress or thermal stress was provided). Also note that some welfare outcomes could also be welfare determinates. For example, it was possible that a study evaluated the effect of mismothering in lamb survival. In this case, behaviour was the determinate and mortality was the outcome. Alternatively, a study might have evaluated the effect of cold weather on mothering behaviour and lamb survival. In this bases, environment was the welfare determinate and behaviour and mortality where the welfare outcomes.

The welfare outcome, abnormal clinical signs included physiological measurements, when a disease was not specified (e.g., tachycardia, dehydration, increased respiration). The abnormal blood/urine, etc. category included abnormal haematology or blood biochemistry measurements, without a specified disease (e.g., elevated cortisol, decreased lymphocytes). Abnormal maternal/lamb behaviour included all impairments to maternal and lamb behaviours, such as lamb stealing, increased time to stand/suckle, and subjective lamb vigour scores and maternal behaviour scores. Abnormal social behaviours included aggression, agonistic behaviours, and bullying. A catch-all category was included to capture all other abnormal behaviours (e.g., vocalizations, hunched back). Health outcomes were categorized by body system, with several specific health outcomes itemized to evaluate their prevalence in the literature (e.g., dystocia, fly strike, lameness). “Other dermatological disease” included exo-parasitism and foot rot, while “Other reproductive disease” included extended duration of parturition, where dystocia was not specifically mentioned. Gastrointestinal disease included internal parasitism. The review team did not resolve conflicts for this outcome, instead providing the broadest mapping of the outcomes.

2.1.1.3. Characteristics of study or study population

The options to report the study population, study design, production system, and farm management system are presented in Table 1.

- The study design: Observational/descriptive studies were defined as observational studies which had no imposed intervention. For example, studies that compared two genetic lines would be observational. Descriptive studies included observational studies without a comparison group i.e., a case series or case reports and studies where an intervention was imposed but no comparison group was included, for example, the vocalization response of a group of animals to castration. Controlled trials and/or experimental studies were defined as studies with two or more comparison groups and a deliberately imposed intervention. Studies that castrated all animals and observed their behaviour would be descriptive, but studies that compared methods of pain mitigation in castrated lambs would be experimental. When the citation did not contain sufficient information for the reviewer to decide upon the design the “not discernible” option was chosen. Only one option could be selected.
- The production system studied: Animals were only considered to be milk producers if a milking machine was used or the animals were described as dairy ewes. Animals were considered to be meat animals if the report included information about carcass quality characteristics or they were described as feedlot sheep. Similarly, wool animals were only reported if they were specifically identified as such by the authors. Inferences were not made from the breed of the animals. If none of these were reported the production system was considered not discernible. Multiple options could be selected.
- The farm management system studied: The management system was considered to be extensive only if explicitly reported such as in outdoor observational studies and experimental studies explicitly evaluating a component of the outdoor environment (e.g., the effect of shelters on the survival of lambs born outside). Similarly intensive management systems were

only reported when explicitly reported in indoor observational studies and experimental studies explicitly evaluating a component of indoor housing (e.g., flooring). The management system was considered not applicable for experimental studies evaluating factors unrelated to indoor and outdoor environments. An option was included to identify studies in which the farm management system was not described. Only one option could be selected.

2.1.1.4. Language of full publication and country of study

The language of the full publication was available within the citation meta-data and was reported by the reviewers as either English, Non-English, or “Can’t tell.” If the country in which the study was located was identified in the abstract, it was reported by the reviewer; “NR” was reported, if the location

Table 1: Data types mapped for citations relevant to sheep welfare

Welfare determinants	Welfare outcomes	Study populations	Study design	Production system	Farm management system
Genetics	Abnormal clinical sign	Undifferentiated flock(s)	Observational/descriptive study	Milk	Extensive
Nutrition/feeding/watering	Abnormal blood/urine, etc.	Ewes	Controlled trial/experimental study	Meat	Intensive
Housing	Abnormal maternal/lamb behaviour	Lambs	Not discernible	Wool	Mixed
Environment	Impaired resting behaviour	Ewes and lambs		Not described	Not applicable - experimental study
Management	Abnormal social behaviour	Rams			Not described
Health	Other abnormal behaviour	Rams, ewes, and lambs			
Behaviour	Death/survival	Wethers			
Handler traits/Human-animal bond	Dystocia	Adults			
	Lameness	Not discernible			
	Fly-strike				
	Impaired/delayed wound healing				
	Other dermatological disease/injury				
	Other musculoskeletal disease/injury				
	Other non-dystocia reproductive disease/injury				
	Poor body condition				
	Birth defects/congenital abnormalities				
	Cardiac disease				
	Gastrointestinal disease/GI parasites				
	Immunological disease				
	Mammary disease				

Welfare determinants	Welfare outcomes	Study populations	Study design	Production system	Farm management system
	Neurological disease Ocular disease Oral disease Respiratory disease Septicemia Urinary disease Other disease/injury Distress Suffering Stress Fear Pain Thermal stress				

2.2. Second-level data characterization

A second level of data characterization was included due to the large amount of data found within the three welfare determinates: management, environment and genetics. Only one of the two reviewer conducted this additional level of screening and this level was added without consultation with EFSA staff. This second-level mapping aided interpretation of the data by further refining the following topic areas. These categories were developed based on the results of the 1st level of screening

For the management-related welfare determinants, a more refined list of management practices was developed and each citation classified based on which was assessed in the abstract. The following exposure categories were used:

- Shearing
- Castration
- Tail docking
- Mulesing
- Isolation
- Weaning/ewe-lamb separation
- Stocking density
- Restraint/handling
- Other

Similarly for the environment-related welfare determinants, after the 1st screening it was clear that some particular environmental topics were more likely to be studied. Therefore the citations were refined into the following exposure categories:

- Thermal stress (hot or cold)
- Light
- Predators
- Noise
- Electricity
- Radiation
- Other

2.3. Relevance screening and data mapping methods

Relevance screening of the titles and abstracts and mapping of the data within the relevant citations were conducted concurrently by two independent reviewers, using the set of questions presented in Appendix B. Due to time constraints at level, conflicts were resolved by a single reviewer.

Inclusion/exclusion conflicts were resolved by the reviewer re-reading the title and abstract. When a conflict occurred regarding either a welfare determinant or a welfare outcome, the responses of the two reviewers were amalgamated. Conflicts regarding study design resulted in a response change to "Not discernible." To resolve other conflicts (e.g., production system, farm management system, population), the reviewer re-read the title and abstract. Relevance screening, data mapping, and conflict resolution were all conducted using the online software DistillerSR⁴. Data were exported to Excel (Microsoft Corp., 2013) for analysis.

3. Identification of relevant questions for systematic review

3.1. Teleconference with EFSA Working Group to discuss D2 deliverable and agree upon a relevant question(s) for systematic review

The D2 deliverable reported the results of the data mapping exercise, and an objective and subjective assessment of the quantity and suitability of the evidence for systematic review. The D2 deliverable was presented to the EFSA Working Group on 24 February, and a teleconference was held to discuss the report and the proposed potential review questions on 4 March 2014.

Following the teleconference discussion, secondary, non-comprehensive, topic-specific, focussed exercises were performed by the review team to clarify the availability of evidence for topics of interest identified by the EFSA Working Group. This exercise was conducted between 4 March and 11 March 2014, for the topics of (1) sheep lameness and (2) sheep body condition score.

A description of the secondary exercises follows.

3.2. Secondary non-comprehensive scoping exercise specific to sheep lameness

3.2.1. Search methods for secondary sheep lameness scoping exercise

The review team at York University conducted a literature search using a single database (Science Citations Index) covering all years and all languages. The search terms used were as follows:

- 1) TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR "hogget*" OR wether*)
- 2) TS=("lame" OR "lameness")
- 3) TS=("footrot*" OR "foot rot" OR "overgrown hoof" OR "overgrown hooves" OR "pododermatitis" OR "podo-dermatitis" OR "scald" OR "digital dermatitis" OR "inter-digital dermatitis" OR "CODD" OR "soil balling" OR "fibroma" OR "granuloma" OR (("foot" OR "feet" OR "toe" OR "toes" OR "white line" OR "whiteline") NEAR/2 "abscess*"))
- 4) 1 AND (2 OR 3)

3.2.2. Relevance screening and data mapping for secondary sheep lameness scoping exercise

References identified by the search were uploaded to DistillerSR®, where they were screened for relevance and data mapped. Due to time constraints, a single reviewer developed the screening and mapping questions, screened references for relevance, and mapped the data of the relevant references. A single question was used to screen the identified references for relevance:

⁴ <https://systematic-review.ca>

- 1) Does the title/abstract describe a study in which a valid estimate of the prevalence of lameness in sheep on commercial (non-experimental) farms may have been reported and the full text is available in English?
 - Yes
 - No
 - Not discernible

References that received a “Yes” response were moved forward for further mapping of their data using the following questions:

- 1) What was being measured?
 - Lameness
 - Footrot
 - Other lameness-related condition (e.g., CODD)
- 2) What best describes the study objective?
 - Primary objective to estimate overall prevalence/incidence in a region unrelated to an outbreak
 - Estimate of prevalence/incidence WAS made secondary to an observational study objective
 - Estimate of prevalence/incidence MAY HAVE BEEN made secondary to an observational study objective (not reported in abstract)
 - Primary objective to estimate prevalence/incidence in a region or purposively sampled flocks related to an outbreak
 - Estimate of prevalence/incidence WAS made secondary to an experimental study objective
 - Estimate of prevalence/incidence MAY HAVE BEEN made secondary to an experimental study objective (not reported in abstract)
- 3) Was there an obvious comparison of prevalence between management systems?
 - Yes
 - No
 - Multiple risk factors evaluated, of which production system may have been one
- 4) Who measured the prevalence?

- Explicitly the researcher
- Explicitly the farmer
- Both researcher and farmer
- Not reported

5) In what country did the study occur?

- Data input by reviewer

3.3. Secondary non-comprehensive scoping exercise specific to sheep body condition score

3.3.1. Search methods for secondary sheep body condition score scoping exercise

The review team at York University conducted a literature search using a single database (Science Citations Index) covering all years and all languages. The search terms used were as follows:

- 1) TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR "hogget*" OR wether*)
- 2) TS=("body condition" OR "body composition" OR "body score*" OR "body scoring" OR "condition score*" OR "condition scoring" OR "fat score*" OR "fat scoring")
- 3) 1 AND 2

3.3.2. Relevance screening and data mapping for secondary sheep lameness scoping exercise

References identified by the search were uploaded to DistillerSR®, the 1st 200 were screened for relevance and data mapped. Due to time constraints, a single reviewer developed the screening and mapping questions, screened references for relevance, and mapped the data of the relevant references. A single question was used to screen the identified references for relevance:

Does the title/abstract describe a study in which a valid estimate of the prevalence of body condition score in sheep on commercial (non-experimental) farms may have been reported and the full text is available in English?

- Yes
- No
- Not discernible

Data mapping was not conducted because of a lack of relevant references in which sheep body condition score was identified as a welfare outcome.

3.4. Follow-up teleconference to discuss findings of the secondary scoping exercises and to agree upon a relevant question(s) for systematic review

A follow-up teleconference was held on 11 March 2014 to discuss the findings of the secondary scoping exercises and to agree upon a relevant question(s) for which a systematic review protocol(s) would be developed.

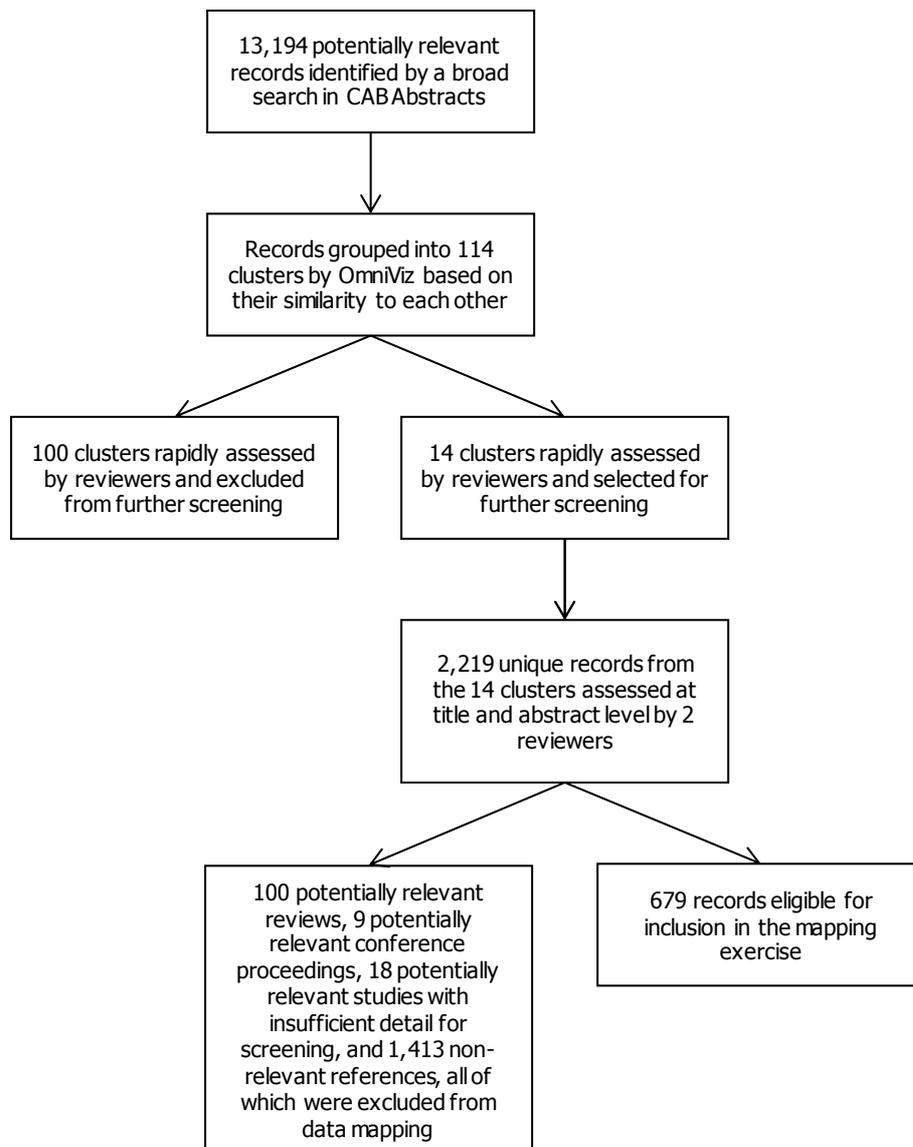
RESULTS

4. Identification of literature for relevance screening and data mapping

The results of the literature identification process for the scoping review are presented in Figure 3. The initial search in CAB Abstracts identified 13,194 citations, which were represented by 114 clusters after text mining. Fourteen clusters were deemed to contain a high percentage of likely relevant citations by the reviewers. These clusters contained 2,219 citations. The results of the cluster analysis for both reviewers were compared and agreement was very high ($Kappa = 0.87$).

Generally, of the terms identified by text mining the citation abstracts, those describing the animal (e.g., “ewe,” “lamb,” “ram”) were found most frequently in the relevant clusters of citations. Some clusters identified as non-relevant also contained these terms (e.g., citations about mating behaviour); however, usually these clusters contained one other non-relevant term (e.g., “sperm,” “spermatozoa”). Other useful terms identified in relevant clusters included “temperament,” “predation,” “preference,” “pain,” “restraint,” “fasting (hunger),” “handling,” “dehydration,” “starvation,” “water deprivation,” and “emotion.” Some terms were found only in non-relevant clusters (e.g., “sward” and “herbage” were found only in non-relevant citations about grazing behaviour/preferences). Unfortunately, many of the most frequently used terms identified were non-specific and, in attempted refinement of the search strategy, returned many irrelevant records (over 17,000 citations were returned), even when combined with the other key elements of the search.

Figure 3: Results of the literature identification process



Because refinement of the search strategy did not result in a literature set that could be screened and mapped within the time constraints of the scoping exercise, the 2,219 citations within the clusters of interest identified by the reviewers were selected as the final body of literature upon which to conduct the scoping exercise.

5. Relevance screening and data mapping

Of the 2,219 citations screened, 679 were identified as relevant to on-farm sheep welfare, 100 were potentially relevant reviews, 9 were potentially relevant conference proceedings i.e. containing multiple conference abstracts, and 18 were potentially relevant studies but lacked sufficient detail in their abstracts to permit full screening (Figure 3).

5.1. Data mapping

5.1.1. Welfare determinants

Over half of the 679 relevant citations (n = 345) described or evaluated a welfare determinant related to management, while a third (n = 215) described environment-related determinants, and almost a quarter (n = 146) described genetic-related determinants (Table 2). When the welfare determinants were further refined in the second-level screening, an overwhelming proportion of relevant studies described the effects of thermal stress on sheep welfare (25% of studies, n = 166) (Table 2). The main management-related welfare determinants described were isolation of individuals, castration, shearing, tail docking, weaning/ewe-lamb separation, and restraint/handling—often a control group for other procedures (Table 3). Examples of the “Other” environment- and management-related welfare determinants reported are presented in Table 3.

Table 2: Welfare determinants identified in 679 citations relevant to on-farm sheep welfare

Welfare determinant	Number of studies (n)
Management	343
Environment	215
Genetics	146
Nutrition/feeding/watering	92
Behaviour	56
Health	53
Housing	39
Handler traits/Human-animal bond	15

* N.B Citations could fall in multiple categories therefore the total number of studies does not sum to 679 citations.

Table 3: Management- and environment-related welfare determinants identified in 679 studies relevant to on-farm sheep welfare

Management-related welfare determinant	Number of studies (n)
Isolation of individuals	76
Restraint/handling	68
Castration	49
Shearing	47
Tail docking	39
Weaning/ewe-lamb separation	32
Mulesing	18
Stocking density	5
Other—separation of group from flock, use of anaesthetics/pain medications/other drugs/ exercise, milking, electro-ejaculation, social mixing, novel environment, ear tagging, restricted suckling, parturition aid, etc.	158
Environment-related welfare determinant	
Thermal stress	166
Light	8
Predators	5
Noise	5
Electricity	5
Radiation	2
Other—dust, terrain, exercise	27

* N.B Citations could fall in multiple categories therefore the total number of studies does not sum to 679 citations.

5.1.2. Welfare outcomes

The main outcome type used to assess welfare impact was the evaluation of some aspect of abnormal or impaired behaviour (i.e., maternal/lamb behaviour, resting behaviour, social behaviour, and other behaviours), with two-thirds of all relevant studies (n = 454) using this method either alone or in combination with other welfare outcomes, such as blood cortisol measurement (Table 4:). Almost half of the relevant studies (n = 315) evaluated a blood, urine, or other sample for changes, the most common test being the evaluation of blood cortisol as an indicator of stress. Abnormal clinical signs were assessed as a welfare outcome in almost a third of relevant studies (n = 214). Changes in body temperature and changes in respiration rate were very common metrics used in this category. Death as a welfare outcome was present in 15% of relevant studies (n = 100); however, health-related outcomes were found infrequently (Table 4:).

Relevant citations that evaluated the impact of genetics on sheep welfare predominately assessed the role of genetics on maternal/lamb behaviour (n = 70 of 146 genetics papers) (Table 5:).

Table 4: Welfare outcomes identified in 679 citations relevant to on-farm sheep welfare

Welfare outcome	Number of studies
Abnormal blood/urine, etc.	315
Abnormal clinical sign	214
Abnormal maternal/lamb behaviour	135
Impaired resting behaviour	65
Abnormal social behaviour	46
Other abnormal behaviour	208
Death	105
Dystocia	28
Lameness	12
Fly-strike	13
Impaired/delayed wound healing	11
Other dermatological disease/injury	8
Other musculoskeletal disease/injury	6
Other non-dystocia reproductive disease/injury	21
Poor body condition	0
Birth defects/congenital abnormalities	4
Cardiac disease	0
Gastrointestinal disease/GI parasites	14
Immunological disease	1
Mammary disease	5
Neurological disease	4
Ocular disease	0
Oral disease	1
Respiratory disease	6
Septicemia	0
Urinary disease	0
Other disease/injury	12
Distress	18
Suffering	0
Stress	39
Fear	20
Pain	35
Thermal stress	38

* N.B Citations could fall in multiple categories therefore the total number of studies does not sum to 679 citations.

Table 5: Welfare outcomes assessed in relevant citations that evaluated the impact of genetics on sheep welfare

Genetics-influenced welfare outcome	Number of citations (n)
Maternal/lamb behaviour	70
Thermal stress	45
Other (e.g., aggression/fear/temperament)	30

* N.B Citations could fall in multiple categories therefore the total number of studies does not sum to 679 citations.

Appendix D contains a table with the frequency count of the welfare outcomes for each welfare determinate category.

5.1.3. Characteristics of the study or study population.

The most frequently studied populations were lambs (n = 261), ewes and lambs (n = 117), and ewes (n = 111), while the most common study design was a controlled trial or experimental study (n = 354) (Table 6). Few citations described the production system or farm management system in which the sheep were studied. Where the farm management system was described, most studies occurred in an extensive management system (n = 89) (Table 6:)

Table 6: Study characteristics of 679 citations relevant to sheep welfare

Population studied	Number of citations (n)
Lambs	261
Ewes and lambs	117
Ewes	111
Rams	35
Wethers	26
Adults (not further specified)	19
Rams, ewes, and lambs	5
Undifferentiated flock(s)	5
Not discernible	100
Study design	
Controlled trial/experimental study	354
Observational/descriptive study	258
Not discernible	67
Production system	
Meat	18
Milk	14
Wool	3
Not described	644
Farm management system	
Extensive	89
Intensive	41
Mixed	10
Not applicable—experimental study	287
Not described—observational study	252

5.1.4. Language of full publication and country of study

The majority of full publications were available in English (n = 576); however, 74 citations were published only in a language other than English (Table 7:). Most citations did not report the country in which the study was conducted (n = 592). In the remaining citations, 35 different countries were

reported. Australia was the most frequently identified country in the citations that reported the country of the study (n = 13) (Table 7:).

Table 7: Language of full publication and country of study in 679 citations relevant to sheep welfare

Language of full publication	Number of citations (n)
English	576
Non-English	74
Can't tell	29
Country of study	
Algeria	1
Argentina	1
Australia	13
Brazil	3
Britain	1
Bulgaria	3
Cameroon	1
Chile	1
China	3
Egypt	8
England	2
Germany	2
Ghana	1
Greece	2
Hungary	1
Iceland	1
India	8
Israel	1
Kazakhstan	1
Netherlands	2
New Zealand	3
Nigeria	1
Northern Ireland	1
Norway	4
Poland	1
Romania	1
Saudi Arabia	2
Scotland	2
Serbia and Montenegro	1
South Africa	4
Spain	1
Sudan	1
Turkey	3
UK	1
USA	5
Not reported	592

6. Identification of relevant questions for systematic review

The teleconference to discuss the findings of the scoping review reported within the D2 deliverable and to agree upon the relevant question(s) for which a protocol(s) would be subsequently drafted was attended by Maria Ferrara, Antonio Velarde, the remaining members of the EFSA Animal Welfare Working Group, Annette O'Connor, Dianna Wolfe, and Hannah Wood. Those discussions are briefly summarized here.

The scoping review for sheep welfare identified significant heterogeneity with respect to welfare determinants, welfare outcomes, and the animal-based metrics used to measure the outcomes. As well as multiple metrics to measure an outcome (e.g., vocalizations, movement, cortisol can all measure pain), the methods to measure most behavioural outcomes (e.g., vocalization) are not standardized and for some pain-related behavioural metrics (e.g., movement), it is unknown if the metric will increase or decrease in response to pain (i.e., sheep may lie down in response to pain or they may become restless). Because of this heterogeneity and lack of standardized metrics, and the resource constraints for a follow-up contract, relevant questions with potential for systematic review were limited.

The review team suggested focusing on a specific mutilation management practice (e.g., castration) and comparing methods (i.e., cutting vs. rubber rings) by measurement of a standard metric (i.e., blood cortisol). The EFSA Working Group's interests lay in the topics of thermal stress (i.e., "What are the extreme temperature values at which sheep become stressed?" "What factors affect the thermal comfort of sheep above or below their comfort zone?"), lamb mortality (e.g., "What are the risk factors for lamb mortality?"), gastrointestinal parasitism, lameness (i.e., "What are the risk factors for lameness?" "What is the prevalence of lameness in different management systems?"), and body condition score (i.e., "What welfare determinants—nutrition, housing, environment, management—impact body condition score?"). The review team advised that the bodies of literature available to review questions related to thermal stress and lamb mortality were too large given the budgetary constraints of the follow-up project; thus, reviews related to thermal stress or lamb mortality would not be feasible in the current frame-work contract. A specific review question and a common outcome measure for gastrointestinal parasitism could not be identified and this topic was not considered further for systematic review. To address the remaining two potential review questions relevant to lameness and body condition score, the review team suggested that limited yet focussed evaluations of the scope of literature could be conducted. These focussed scoping exercises would search only a single literature database (CABI), thereby limiting the number of citations necessary to screen for relevance, given the short 7-day timeframe, but still providing an adequate general understanding of the breadth and scope of literature available for both topics. It was made clear by the review team that this exercise would not be considered a comprehensive scoping and that a more comprehensive search of multiple databases would need to be conducted in the follow-up project, should a topic be identified for systematic review. The teleconference was adjourned so that the Working Group could discuss amongst themselves the questions of interest to them, within the topics of sheep lameness and body condition score. A follow-up email from Maria Ferrara to Annette O'Connor on 5 March 2014 identified the following questions of interest to the Working Group:

- 1) What is the prevalence of lameness in sheep and its severity in the described farming systems?
- 2) What is the prevalence and, where indicated, the temporal (seasonal) variation of poor body condition in sheep in the described farming system?

The focussed scoping exercises proceeded and a teleconference was scheduled the following week to discuss the findings.

6.1. Secondary non-comprehensive scoping exercise specific to sheep lameness

762 references were identified in the search and available for relevance screening. Thirty-four references were deemed relevant to the screening criteria and thought to potentially provide an estimate of the prevalence of lameness or a lameness-related condition. One reference compared the prevalence of lameness by management system type; however, the study from Iran was located outside of the European region, as defined by the OIE. Eight references studied the impact of multiple

risk factors on the prevalence of lameness or a lameness-related condition but didn't report management system as one of the risk factors evaluated. Review of the full text would be necessary to determine if the effect of management system had been assessed. Six of these eight studies were conducted in locations within the European region. There was heterogeneity of the outcomes assessed within these 6 studies, with lameness, footrot, and other lameness-related conditions occurring.

6.2. Secondary non-comprehensive scoping exercise for sheep body condition score

A total of 1,757 references were identified in the non-comprehensive search and available for relevance screening. The majority of literature pertaining to body condition score reported body condition score as an independent variable (i.e., assessing the effects of body condition score on a health or production outcome such as fecundity) and not an outcome; thus, it was not relevant to the question of interest posed by the EFSA Working Group. Because of this paucity of relevant literature, no further mapping of data was undertaken for this topic.

6.3. Follow-up teleconference to discuss findings of the secondary scoping exercises and to agree upon a relevant question(s) for systematic review

The teleconference subsequent to the focussed scoping exercises to agree upon a final relevant question(s) (if any) for which a protocol(s) would be subsequently drafted was attended by Maria Ferrara, Antonio Velarde, Annette O'Connor, and Dianna Wolfe.

The scoping exercise that focussed on literature related to body condition score as a welfare outcome identified minimal literature pertinent to the topic. Given the paucity of literature available on the topic, body condition score as a welfare outcome was not considered further for systematic review.

The findings of the scoping exercise that focussed on literature related to lameness suggested that a systematic review protocol could be developed to evaluate the effect of management system type on the prevalence of lameness and lameness-related conditions, with further refinement of a PECOS question. In further discussion, it was decided to constrain a review to studies located within the European region, as defined by the OIE. As well, the review team advised that because season of the year may be associated with both management system and the prevalence of lameness/footrot/etc. (i.e., a confounder), seasonal data should be extracted from the relevant papers, when reported. The EFSA team suggested that breed-related differences in the prevalence of lameness may also be of interest. The review team agreed that breed-related data would be extracted, when reported in relevant papers that evaluated management system as a risk factor for lameness. The review team would not conduct a separate systematic review evaluating the impact of breed on lameness and lameness-related conditions.

The outcome of the final consultations after the scoping review was conducted was that the review team would develop a protocol for the systematic review of literature evaluating the impact of management system type on the prevalence of lameness and lameness-related conditions. The protocol developed (Appendix I) and the results of that review are presented in the subsequent section of this document.

CONCLUSIONS

The results of the scoping exercise, clarified the topics that were studied in the dataset identified. The tables, MS Excel spread sheets and Endnote libraries provided to EFSA will provide the end-users with a summary of the topics studied. The scoping review however does not provide the findings of the studies. The end-users are able to use the materials provided to rapidly identify groups of citations that focus on welfare determinants and outcomes of interest. For example, if the end-users are interested in evaluating the associations between "non-dystocia reproductive outcomes" and genetics,

the review identified 13 studies with that combination. These studies can easily be retrieved and the associations evaluated. This is likely preferable to conducting a search directly in a citation database such as CABI because the experience of this project shows that a very large number of non-relevant citations would need to be screened to identify relevant studies.

It is of course of interest to know if the subsets identified by the screening process are representative of all studies on the topic. This is a question of the external validity of the initial screening set. Such issues are judgements. A factor in favour of concluding the literature is representative is that the original search used welfare terms that were non-specific; therefore, a wide range of topics were identified. Further, the search was conducted in CABI, which is one of the most comprehensive databases for veterinary topics. However, good coverage of veterinary topics is not direct evidence of good coverage of welfare topics. We are unaware of any assessments of coverage of citation bases for welfare topics. A factor against concluding the literature is representative was the use of only a subset of the original 12,000 references identified by the search. Time limitations prevented screening of all citations and without examination of the “other” clusters, we cannot be sure that important topics were excluded. Another issue is that clearly the scope of welfare topics includes every aspect of production; however, many production-related studies may cover relevant welfare issues without including the terms “stress,” “welfare,” or “behaviour.” Without reviewing all the papers about sheep production, it is unclear how a scoping review could truly identify all papers relevant to welfare.

The topics proposed by the review team for subsequent systematic reviews, which focused on management practices and frequently reported outcomes such as cortisol, were not of interest to the EFSA working group. Instead the EFSA working group proposed to focus on lameness and body condition score. These topics did not feature frequently in the scoping review. New searches were designed and new screening exercises conducted to evaluate the types of studies present. Consistent with the findings of the scoping review, the searches conducted did not identify a large amount of relevant literature about the welfare determinants of these two outcomes. Large bodies of literature exist about the pathophysiology and treatment of lameness, and about the impacts of body condition on reproduction and carcass characteristics; however, these are not relevant to the review topics proposed. Included in the contract is a protocol designed to assess the association between the prevalence of lameness and management system (outdoors or indoors). Given that management system is an exposure that cannot be assigned randomly, the review will use observational studies to evaluate this association. We anticipated that few relevant studies will be identified. As well, given that the outcome of interest is measured at the group level and confounding is a common issue in observational studies, the design and statistical methods used in any identified relevant studies will be complex.

A SYSTEMATIC REVIEW OF THE EFFECT OF EXTENSIVE/OUTDOOR/MIGRATORY MANAGEMENT ON LAMENESS COMPARED TO INTENSIVE/INDOOR MANAGEMENT SYSTEMS IN SHEEP RAISED FOR THE PRODUCTION OF MEAT, MILK, OR WOOL IN EUROPE

INTRODUCTION AND OBJECTIVES

The specific overall purpose of this aspect of the project was to conduct of systematic review that would compare lameness in extensive/outdoor/migratory management compared to lameness in intensive/indoor management systems. This topic was chosen after a scoping review was conducted to identify the main welfare risks related to the farming of sheep for wool, meat and milk production. The results of the scoping review are available from EFSA (start-up contract- Call Reference: RC/EFSA/AHAW/2013/03).

MATERIALS AND METHODS

6.4. PECO(S) Question

The review question was “In sheep raised for the production of meat, milk, or wool on commercial farms located in Europe (defined as member countries of the OIE European Region) (P), what is the impact of extensive/outdoor/migratory management (E) compared to intensive/indoor management (C) on the farm-level prevalence or incidence of lameness, footrot, and other lameness-related conditions measured during non-outbreak periods (O) in observational studies or experimental studies with management type as the exposure of interest allocated at the farm-level (S)?”

6.5. Protocol and registration.

The protocol for this review was developed as part of the contract EFSA-AWAH-2013-3 and is included in Appendix Appendix I. .

Several deviations from the protocol occurred. The protocol originally indicated that relevant studies would only include group level measures of the outcome based on the rationale that the exposure would occur at the group level and not vary within farm. During screening of full texts we identified one study conducted on multiple farms that used an animal level analysis. Given the study was conducted on multiple farms results of the paper were otherwise relevant we deviated from the protocol to include this study. A second deviation from the protocol was to report results from studies that assessed the association between stocking density and lameness and housing and lameness. The rationale for including these exposures was that definitions of “outdoor management systems” are imprecise and it was therefore more comprehensive to include such information. This decision was made prior to screening of relevant studies.

6.6. Eligibility criteria

6.6.1. Relevant populations

The relevant study population was defined as sheep raised for the commercial production of meat, milk, or wool in the European region. The European region was defined as member countries of the OIE European region (http://www.oie.int/fileadmin/Home/eng/About_us/docs/pdf/2009_Commission_Europe_A.pdf), and includes northern Asia (e.g., Russia, Ukraine, Uzbekistan), Turkey, Israel, and several other non-EU countries.

6.6.2. Relevant exposure or comparisons

Relevant studies must have reported the assessment of the impact of the management system type under which the sheep were raised (e.g., outdoors, indoors, or mixed (both indoors and outdoors)). Management system was defined as the type of housing as it related to exposure to the environment (i.e., indoors or outdoors). Outdoor housing was also not defined as an absence of structures. For example, feedlot animals may be raised in an outdoor management system but confined in pens and animals in extensive outdoor grazing systems may be offered shade structures. These latter two examples were still defined as outdoor systems. Systems that house animals for management practices, such as housing for lambing or shearing, but otherwise place animals on pasture were defined as outdoor systems. Finally, outdoor management was not limited to meaning access to pasture, as feedlot animals may not have access to pasture but are raised outdoors. Studies that reported the association between stocking density and lameness were considered relevant. Such studies had to explicitly report a density of animals per unit of space.

6.6.3. Types of outcome measures

Outcomes of interest were measures of lameness and lameness-related conditions. Lameness-related conditions included (but were not limited to) footrot, contagious ovine digital dermatitis, interdigital dermatitis, scald, white-line disease.

6.6.4. Relevant study designs

Observational studies relevant to the review were cohort, case-control, or cross-sectional studies that measured the association between management types (e.g., extensive/outdoor/migratory and intensive/indoor systems) and incidence or prevalence of lameness, footrot, or another lameness-related condition. Studies measured at the farm level were required to have more than one observation for exposures of interest (i.e., more than one farm per exposure level). Observational studies that compared the prevalence of lameness on one farm with a single type of management system with the prevalence of lameness on another single farm with a different type of management system were not included due to confounding by other practices between farms. Although the association between the outcome and the exposure of interest could be confounded by other factors, we did not exclude studies that only conducted univariable analyses.

6.7. Information Sources

A range of information sources indexing published research were searched for studies reporting on lameness in sheep (Table 8:)

Information on on-going or recently completed trials, unpublished research, and research reported in the grey literature was identified by searching databases that index conference proceedings and specialised search engines:

- Conference Proceedings Citation Index – Science
- Science.gov
- ScienceResearch.com
- Open Grey

The reference lists of relevant studies were hand searched for additional potentially relevant citations.

6.8. Search strategy

6.8.1. Search terms

The search terms consisted of two key elements that are presented for Science Citation Index (Web of Knowledge; Thompson Reuters) in Figure 1:

- The population: sheep (search line 1)
- The exposure: lameness (search lines 2 to 8)

The search terms for the exposure element included common causes of lameness in farmed sheep such as foot-rot, contagious ovine digital dermatitis, and shelly hoof. An expert in feet diseases in sheep was consulted by EFSA to ensure that the diseases included in the search were appropriate and that there were no significant omissions.

To ensure the search was adequately sensitive, the concept of prevalence was not built into the search strategy. This concept is often poorly or inconsistently described in the title, abstract, and indexing of database records and, therefore, its inclusion may have resulted in relevant studies being missed.

The searches were not limited by date, language, or study design. However, where the functionality of the database interface allowed and the volume of records retrieved was high, the searches were limited to attempt to restrict the results to those related to a European context. Studies indexed with countries which do not belong to the OIE Regional Commission for Europe were excluded using the NOT operator.

The search strategy developed for the Science Citation Index was adapted appropriately to perform efficiently in other information sources. Adaption included a consideration of database interface differences (search syntax) as well as adaption to different indexing languages. The strategies used to search each information source are presented in Appendix A.

6.8.2. Search considerations

Given the short timelines for the review from the contractor (8 weeks), translations were not conducted for papers that were not available in English. We included papers that could be obtained within 1 month after the start of the contract. Potentially relevant papers identified but not obtained within that time frame have been indicated in the Results section of this report.

6.8.3. Search results

The results of the searches were downloaded into bibliographic management software (EndNote 7) and de-duplicated using several algorithms. All remaining references were uploaded to DistillerSR® (Evidence Partners©, Canada, 2012), an internet-based systematic review software, for relevance screening, data extraction, and data management. Within DistillerSR®, de-duplication was repeated. To remove references related to foot-and-mouth disease, searches were conducted in DistillerSR® for “foot-and-mouth” and “FMD,” and the identified references were quarantined (i.e., removed from the list of references available for relevance screening).

6.9. Study Selection

Two reviewers independently performed the relevance screening exercise, with disagreements on the inclusion or exclusion of studies resolved by consensus or with the input of the review coordinator.

6.9.1. Screening for eligibility for the review

Two levels of screening were used to identify relevant studies. The 1st level was based on title and abstract and identified studies that assessed risk factors for lameness. The second level of screening was based on the full text. The questions used to determine inclusion in the review are reported in Appendix B. The list of other independent variables that had been evaluated for association with lameness or a lameness-related condition was also extracted from the full text.

6.10. Data collection process

Data extraction forms were designed in DistillerSR®. Initial forms were designed and piloted on several papers and modified as required for use as discussion. Two reviewers independently extracted data from studies deemed to be relevant to the review and conflicts were resolved by the senior reviewer.

6.11. Data Items

6.11.1. General study characteristics

The following data were extracted from each relevant manuscript: study design, year the study was performed, country, sheep production type, class of sheep studied for the estimate of lameness, breed of sheep enrolled, time of year in which the data was collected, sampling method of farms/flocks, total number of farms/flocks enrolled in the study, sampling method of animals within farms/flocks and total number of animals enrolled in study. The exact form used and notes used by data extractors to interpret the variables are reported in Appendix F.

6.11.2. Outcome-level and exposure information

If more than one outcome was evaluated in a study, separate outcome-level information was extracted for each outcome. For each outcome we extracted: the condition assessed, the unit of analysis for the outcome, the authors definition of the outcome copied directly from the full paper, outcome type, outcome measurement methods. For continuous outcomes the central tendency metric and the dispersion metric(s) reported for each exposure group were extracted for each outcome. For categorical outcomes attempts were made to extract, the number of units with the outcome (r) and the number enrolled (n) in each exposure group.

For the association between the exposure and the outcome, we extracted the effect size measure, the precision measures reported for the effect size and the statistical methods used to calculate the effect size. When extracting the outcome, the measures of association were transformed to ensure consistent referent category. If only measure of effect and p value was reported for a univariable or multivariable model the confidence interval was back calculated making the assumption that the p was derived from a Wald test for the regression coefficient using a Z distribution for the p value. Regardless of the original referent, no grazing or no housing was used as the referent For multivariable methods, covariates included in the model were extracted. The exact form used and notes used by data extractors to interpret the form questions are reported in Appendix G.

6.11.3. Exposure information

For each exposure we extracted the authors definition of the management exposure. The exact form used and notes used by data extractors to interpret the form questions are reported in Appendix G.

6.12. Assessment of risk bias in included studies

Several tools are available for the assessment of the risk of bias for observational studies. Our preference was to use a tool for non-randomized studies from Cochran Group. However, that tool was being tested and modified at the time this review was conducted (personnel communication with Cochran Risk of Bias group). Therefore we used the risk of bias tool and confounding for observational studies of interventions or exposures, developed by the Agency for Healthcare Research and Quality (<http://www.ncbi.nlm.nih.gov/books/NBK154460/>), as required to be suitable for the studies evaluated in this review. In addition to the items recommended, space was available for reviewer to comment on other sources of bias. The exact form used and link to the document used by data extractors to interpret the form questions are reported in Appendix Appendix H.

6.13. Summary measures

Studies reported either prevalence ratios, odds ratios, or rate ratios. The log of the measure of association and standard error used for graphing were back calculated from the extracted point estimates and 95% confidence intervals using RevMan(RevMan, 2012). When authors conducted a multivariable analysis the adjusted measure of association was reported in preference to the unadjusted. However when only an unadjusted estimate was available, this was extracted, reported and used in any analyse.

6.14. Data synthesis

A forest plot was created to display the data for both exposure variables. Variables that were related to management system (pasture access) were grouped together. As the exposure categories were not truly equivalent across the studies and some animals were used for multiple measures a summary effect size was not calculated. It was not possible to conduct statistical tests to assess if clinical or methodological factors might be associated heterogeneity because insufficient independent studies were available. Given the short timelines for the contract, we did not plan to contact authors to obtain missing data.

6.15. Data synthesis

It was not possible to assess small study effects for the following reasons: insufficient studies measuring the same outcome, exposure categories were not truly equivalent across the studies and some animals were used for multiple measures.

6.16. Additional Analyses

No additional unplanned analyses were conducted.

RESULTS

6.17. Search results

The searches yielded 7,007 records. The source of these records is presented in Table 9: . After duplicates were removed, 3,991 records remained, which were uploaded to DistillerSR®. Thirty-one additional references were identified as duplicates and 742 references were identified as containing either “foot-and-mouth” or “FMD.” These 783 references were removed, leaving 3,208 references available for relevance screening. A review of the references of the papers included in the final review did not reveal any further potentially relevant citations. The flow of references from the search results to synthesis is presented in Figure 2.

6.18. Study selection

6.18.1. Screening for eligibility for the review

Twenty-two of the 3,208 references available (0.7%) were identified as potentially relevant in the first level of relevance screening, in which titles and abstracts were evaluated. The full text of one of the 22 references could not be obtained within 1 month of the contract start date. Of the 21 references for which full texts could be obtained, 6 were identified as relevant to the review and were moved forward for data extraction.

Nine of the 21 full texts available for the second level of screening reported risk factor analyses for lameness or a lameness-related condition. All other independent variables (i.e., excluding management system and stocking density) that had been evaluated for association with a lameness outcome are summarized in Table 10.

6.19. Study characteristics

Two studies had missing information for the meta-analysis. For one study we back calculated the confidence intervals from the odds ratio and p value (Gelasakis et al., 2013). Another study did not report any data for an outcome of interest. The authors reported assessing housing but no results were provided and the categories assessed were not reported (Angell et al., 2014).

6.19.1. General study characteristics (potential sources of clinical heterogeneity)

The general characteristics of the relevant studies with management system or stocking density as an exposure are summarized in Table 11: Four of the relevant studies were postal surveys of sheep producers in England, or England and Wales. One author's name (L.E. Green) was common to all 3 of these papers.

Only one study, the longitudinal study in Greece, reported the production type of sheep studied (milk). This study and one of the postal surveys assessed only ewes, while the other postal surveys evaluated the entire flock. The Greek study and one of the postal surveys used some form of stratified sampling to obtain their study flocks, while the two other postal surveys used a convenience sampling method. All studies conducted a census of all animals (or ewes) in the study flocks as their animal-level sampling strategy.

6.19.2. Exposure level information

The definitions of the exposures evaluated are provided in Table 12: ,Table 13: and Table 14: No studies explicitly used the terms extensive/outdoor/migratory management or intensive/indoor management. The review team therefore interpreted the authors variable descriptions and included those that likely compared the systems of management. When studies compared animals based on either access to pasture or the amount of time spent indoors we summarized these as management systems. There is heterogeneity in this body of work, the exposure group (column two of Table 13:) sometimes contains sheep with no access to pasture (Gelasakis et al., 2013) and some animals confined just for winter (Wassink et al., 2004). Similarly, the referent group (column three of Table 13:) includes some animals that are entirely pasture based (Gelasakis et al., 2013) and others that may be housed for up to 3 months (Green et al., 2004). The reviewers did make the assumption that no winter housing equates to no housing, i.e., that animals are rarely housed in spring, summer or autumn. However no studies stated this explicitly.

Some studies reported stocking density, which does not equate to indoor or outdoor management, however it was decided to extract and report these data as there is the assumption that higher stocking

densities may be associated with indoor housing. Only two measures of stocking density were studied because only two research groups included this question Table 14: One concern is that the rationale for the cut-offs for the density or duration of housing are not provided by any authors and it is unclear if the “right” threshold for densities or time housed were studied.

6.19.3. Outcome-level information

6.19.3.1. Management system - extensive/outdoor/migratory management (E) or intensive/indoor management

The approach to measuring lameness varied by manuscript. Three of the relevant studies with management system exposures had multiple outcomes—two studies had 2 outcomes each and another had 3 outcomes—meaning there were 9 outcomes assessed within the 5 studies. The outcome-level data are presented in Table 12: .

Of the 9 outcomes evaluated, 3 were related to lameness, 3 were related to only interdigital dermatitis, 2 were related to only footrot, and 1 was related to the occurrence of either footrot or interdigital dermatitis. Three of the outcomes were evaluated at the animal level, in the 2 longitudinal studies (i.e., the studies that were not postal surveys). The variety of study designs, classes of sheep studied, levels of analysis, and outcome types led to considerable heterogeneity in outcome definitions.

6.19.3.2. Stocking density

Four references evaluated the association between stocking density and lameness (Table 14: One of the four references had 3 outcomes of interest, two references had 2 outcomes of interest, and the remaining study had one outcome of interest. Thus, there were 8 outcomes assessed within the 4 stocking density studies.

6.20. Assessment of risk of bias within studies

The risk of bias assessment is provided in Table 15. Concerns about selection bias were not major. With respect to selection bias, the major concern would enrolment of exposed group with respect to a particular outcome or an outcome with respect to an exposure. This seems unlikely in the studies reported here. Selection bias occurs when one group differentially enrolls. Although it is possible that producers with lameness were more likely to enrol, it seems unlikely, given the number of questions asked in most of the surveys, that producers with lameness and a particular management system or stocking density would have been more inclined to return the survey. So the risk of bias from selection bias seems low.

Misclassification of the exposure or the outcome was a concern as a source of bias. None of the studies used an approach to assessing the outcome that was blind to the exposure. The concern is that the associations observed are a function of differential sensitivity and specificity of lameness diagnosis. For the variables of interest, if no association exists it would have been necessary for producers with limited access to pasture and higher stocking density to overestimate lameness and those without those characteristics to underestimate lameness. Clearly this cannot be determined empirically, however if these factors had been reported as risk factors, it is possible that the bias would be away from the null. One study did report advice producers were given to prevent foot-rot and it did not include information about stocking density or access to pasture. In the senior reviewers opinion (AOC) it seems more likely that the estimates of lameness would be affected by how recently a producer had a lameness outcome rather than a management characteristic. The risk of bias from selection bias is unclear. One group of researchers did provide data suggesting that producers can accurately detect the diseases when presented with concurrent cases. Regrettably this does not

translate to validation that producers can accurately recall disease occurrence or accurately estimate prevalence.

Only one study explicitly reported assessment of confounding (Table 15). Other studies may have adjusted for confounding by inclusion of covariates but formal assessment of confounding was not conducted. The two factors studied are likely confounders. For example, animals housed indoors tend to have higher stocking density than outdoor animals, and it is therefore hard to separate the two factors. Given the number of factors that may be associated with grazing and lameness, the failure to adjust for confounding in many of the studies suggests that the risk of bias from confounding bias is high.

The two studies conducted at the individual animal level did not appear to adjust for the farm effect. If this interpretation is correct, these studies are likely to underestimate the variation and be more likely to report a statistically significant association.

6.21. Results of risk of bias within studies

The results for studies that assessed exposure to pasture and lameness are reported in Table 13: and Figure 3. Overall, the studies suggest either no association or an increased lameness in animals that spend more time indoors. The studies that report odds ratios show stronger associations based on the point estimates. However this is potentially misleading, as all but one study report a confidence interval that includes one. Also the odds ratio is always further from the null than the risk ratio and the difference between the risk (prevalence) ratio and odds ratio is larger when the disease is common as occurred in many of these studies. As no studies reported the baseline prevalence of risk and some estimates were from adjusted models. However, given how common lameness was in the studies it seem likely that the distance from the odds ratio and risk ratio estimates would be large.

Only one study reported a strongly protective association in a univariable model. These authors reported that “*Sheep which had experienced their highest prevalence of footrot between December 1999 and April 2000 and which were not housed had a significantly higher prevalence footrot than housed sheep which had experienced their highest prevalence of footrot during their housing period (15.1 v 7-6 cases of footrot per 100 ewes per day)*”(Wassink et al., 2003). However, after adjustment for other covariates this exposure was apparently no longer significantly associated with the outcome. The adjusted association and confidence interval were not reported.

One study reported a strong risk factor association (odds ratio = 5.52) that was not consistent with other studies (Green et al., 2004). The study is reported in a conference proceeding and some details about the approach to analysis are unclear. For example, the model is reported as having an additive error in a logistic model, observations appear to be assumed as independent and no apparent adjustment for confounders is reported. These factors may not be important but without a clearer description of the model it is unclear if they have any impact on the reported strong association. Further, this is the only study that has as a referent a group of animals that may actually have some housing. This study compared animals with “housing from 3 to 6 months” to animals with “housing for < 3 months”. This means that this comparison may represent a comparison of long term housing to short-term housing. It may also represent a comparison of long-term housing with no housing. Finally it also could be a comparison of long-term housing with combined short-term and no housing. All other studies had no housing as the referent. Therefore it is unclear if this comparison group is valid and if it should be removed from the others. The data presented in Figure 4 exclude this study so that the comparison groups are more similar and less ambiguous.

The results for studies that assessed stocking density and lameness are reported in Table 14: and Figure 5. Although few studies have evaluated this outcome, the finding was reasonably consistent that higher stocking density was associated with more lameness.

6.21.1. Risk of bias across studies

The consistent issues across the studies include poor outcome measures and lack of assessment of potential confounders. There is also evidence of selective reporting. Some studies reported the magnitude of effect (and precision) when statistically significant in univariable models but then failed to report the same information when not statistically significant in multivariable models. This approach to reporting can create a false impression in meta-analyses where there is an emphasis on extracting the magnitude of effect, because non-significant associations common in multivariable models essentially become missing data. Selective reporting bias is a form of publication bias. We were not able to evaluate small study effects i.e., if larger findings were associated with smaller studies, as insufficient data were available.

6.22. Summary of evidence

The reviewers assessment is that overall this is a weak body of work which provided evidence that management system (i.e., access to pastures) is not associated with more lameness. With respect to stocking density, this weak body of work suggests there is evidence that higher stocking densities are associated with more lameness. Unfortunately it was not possible to quantify the stocking densities that are likely to cause issues or to quantify the magnitude of the effect i.e., it was not possible to report that “at stocking density (x ewes per m²) the risk of lameness increased by 50%” because of the heterogeneity in exposures and outcomes.

6.23. Limitations

Common sources of bias in observational studies are selection, information and confounding bias. It should be noted that many of the studies are quite large and this does add strength to the body of work. Often in veterinary science very few farms are recruited but in this area the researchers have often attempted to conduct large representative postal surveys. This is reflected in comments that selection bias is of minimal concern in this body of work overall. Authors have also attempted to improve the measurement of the outcome lameness. Several studies reported validating that farmers can recognize the disease correctly. However, no studies reported that farmers can correctly recall the incidence of lameness. Such validation would require comparison of farmer recalled prevalence with records made at the time. No studies reported using computerized records for estimates of prevalence. Finally, failure to adjust for or report confounding variables means that the research community is not aware of the potential confounders of the association between lameness and management systems. When reporting associations authors should report those observed in multivariable models even when not statistically significant. This can be problematic when model-building approach's remove variables. Options include reporting the association when the variable is removed from the model, or creating models with all variables in the model. This would limit the impact of selective reporting. Many studies would benefit from following reporting guidelines for observational studies (von Elm et al., 2007).

CONCLUSIONS

In the reviewers assessment the body of work provides evidence that management system is not associated with more lameness and higher stocking densities are associated with more lameness. The body of work may appear to be quite small however given the difficulties faced by researchers

investigating this topic identifying 6 studies that looked at these factors in a limited region of the world is a reasonable body of work. This is a difficult topic to study as the exposures are variable and the outcome difficult to measure in production systems as it can occur year round and has numerous causes. Like many judgements about disease causation this assessment is based on currently available information. It should not come as a surprise that if researchers were able to address the concerns about measurement of the outcome and adjustment for confounding this conclusion may be adjusted.

Table 8: Information sources searched to identify relevant studies

Database	Interface
Science Citation Index (SCI)	Web of Knowledge, Thompson Reuters
Conference Proceedings Citation Index – Science (CPCI-S)	Web of Knowledge, Thompson Reuters
CAB Abstracts	Web of Knowledge, Thompson Reuters
BIOSIS Citation Index	Web of Knowledge, Thompson Reuters
MEDLINE and MEDLINE In-Process	Ovid
AGRIS	http://agris.fao.org/
AGRICOLA	http://agricola.nal.usda.gov/
TEKTRAN	www.ars.usda.gov/services/tektran.htm
CRIS	http://cris.nifa.usda.gov/

Table 9: Number of records retrieved by information source

Information source	Number of records identified
Science Citation Index (SCI)	1,377
Conference Proceedings Citation Index – Science (CPCI-S)	68
CAB Abstracts	2,019
BIOSIS Citation Index	1,500
MEDLINE and MEDLINE In-Process	1,202
AGRIS	466
AGRICOLA	299
TEKTRAN	0
CRIS	0
Science.gov	0
ScienceResearch.com	73
Open Grey	3
Total	7,007
Total without duplicates	3,991

Table 10: Independent variables (except management system) evaluated for association with lameness or a lameness-related condition in risk factor analyses that were potentially relevant in a systematic review of the effects of management system on lameness or lameness-related conditions in sheep in Europe

Independent variable grouping	Examples of independent variables in grouping	Number of studies
Foot trimming	No routine trimming vs. Once or more/year Trim sheep with footrot (y/n)	5
Treatment/medication for foot lesions	Antibiotic injections (y/n) Antibiotic sprays (y/n) Zinc sulphate products vs. Formaldehyde-based products Identification of treated sheep (y/n)	4
Farm topography	Minimum elevation above sea level Maximum elevation above sea level Average elevation Hill/mountain vs. Upland vs. Lowland	4
Flock size	Number of ewes on the farm Total number of breeding ewes Total number of sheep	3
Vaccination	Do not vaccinate vs. Vaccinate ewes or rams only vs. Ewes and rams only vs. Other	3
Foot bathing	Routine treatment (y/n) Individual treatment (y/n) Standing surface after foot bathing	3
Catch/isolation of affected sheep	Separate lame sheep at pasture (y/n) Separate lame sheep at housing (y/n) Individual treatment (y/n) Catch lame sheep (y/n) Isolate sheep with footrot (y/n)	3
Presence of concurrent disease	Prevalence of other foot conditions in flock Number of animals with condition in previous 4 weeks	3
Housing	Lambing inside (y/n) Litter depth (Warm, damp, and thick vs. Thin and dry)	2
Isolation/quarantine of purchased animals	Some or none vs. All	2
Farm type/purpose	Pedigree/replacement/terminal sires (y/n) Meat lambs/store lambs (y/n)	2

	Hobby farm (y/n)	
	Conventional vs. Organic	
Culling for footrot	(y/n)	2
Breed	Bergamasca vs. Sarda vs. Massese vs. Padovana	1
Other animal-level factor	Parity	1
Direct contact with positive sheep	Purchase sheep from positive flock (y/n)	1
	Contact with positive flock that has trespassed fences (y/n)	
	Contact with positive flock after trespassing fences (y/n)	
	Borrowed ram from positive flock (y/n)	
	Shared pasture with another flock (y/n)	
Indirect contact with positive sheep	Shared weighing scales with positive flock (y/n)	1
	Shared road to pasture with positive flock (y/n)	
Neighbouring flocks positive for footrot	At least one positive flock within 0–1 km (y/n)	1
	At least one positive flock within 1–2 km (y/n)	
	At least one positive flock within 2–3 km (y/n)	
Farm acreage	5-level categorical variable	1
Time of year	Early vs. Late lambing	1
Presence of cattle on farm	(y/n)	1
Showing animals	(y/n)	1

Table 11: General characteristics of relevant studies with management system as an exposure identified in a systematic review of the effects of management system on lameness or lameness-related conditions in sheep in Europe

Citation	Design	Year	Country	Production type	Class of sheep	Breed	Flock sampling method	Number of flocks	Animal sampling method	Number of animals	Time of year
(Angell et al., 2014)	Cross-sectional	2012	Wales	NR	Flocks	Welsh Mountain Sheep, Texel, Mule, Suffolk, and others not described	Random sample from sampling frame	511	Census of all animals on farm	Not reported	Not discernible
(Gelasakis et al., 2013).	Other	2008	Greece	Milk	Ewes	Chios-breed dairy	Not described	7	Census of all animals on farm	166	January, to July
(Green et al., 2004).	Other	2002	UK	NR	Not described	Not described	Random sample after k-means cluster analysis	9	Census of all animals on farm	1,618	Full year
(Kaler and Green, 2009)	Cross-sectional	2005	England	NR	Flocks	Texel, Mule, Suffolk, Swaledale, Polled Dorset, Lley, Beltex	Stratified	809 dataset A/ 443 dataset B	Census of all animals on farm	NR	Full year
(Wassink et al., 2003).	Cross-sectional	2000	England and Wales	NR	Ewes	Not described	Convenience	209	Census of all animals on farm	67,514	Full year
(Wassink et al., 2004)	Cross-sectional	2000	England and Wales	NR	Flocks	Not described	Convenience	209	Census of all animals	NR	Full year

Table 12: Outcome and exposure definitions of studies identified as relevant in a systematic review of the effects of management system on lameness or lameness-related conditions in sheep in Europe

Citation	Outcome definition	Temporality of measurement relative to disease occurrence	Person who measured outcome	Exposure(s) measured
(Angell et al., 2014).	Reported presence of contagious ovine digital dermatitis in flock		Farmer	Housing ¹
(Gelasakis et al., 2013).	Lameness occurrence during study period: y/n	Sheep observed every 2 weeks	Researcher	Management system Stocking density
(Green et al., 2004).	New occurrence of footrot or interdigital dermatitis during a 14-day interval during the study period. A positive outcome could occur more than once during the study period.	Prospectively collected data over 15 months	Farmer	Management system
(Kaler and Green, 2009).	Prevalence of lameness in flock in 2004	Recall in 2005	Farmer	Management system Stocking density
(Kaler and Green, 2009).	Prevalence of interdigital dermatitis in flock in 2004	Recall in 2005	Farmer	Management system Stocking density
(Kaler and Green, 2009).	Prevalence of footrot in flock in 2004	Recall in 2005	Farmer	Management system Stocking density
(Wassink et al., 2003)	Daily number of ewes with footrot in the worst affected months during the study period. This was a subset analysis of only flocks that had highest prevalence of footrot from Dec-Mar.	Recall in 2005	Farmer	Management system Stocking density
(Wassink et al., 2004)	Mean prevalence of interdigital dermatitis of 5% or more of the ewe flock (y/n) during months with most cases	Recall in 2005	Farmer	Management system Stocking density
(Wassink et al., 2004)	Mean prevalence of interdigital dermatitis of 5% or more of the lambs (y/n) during months with most cases	Recall in 2005	Farmer	Management system Stocking density

¹ Study reported a housing variable that may have included pasture access but no details provided or results

Table 13: Outcome data summarized by exposure levels for relevant studies identified in a systematic review of the effects of management system on lameness or lameness-related conditions in sheep in Europe.

Citation	Exposed group (less or no access to pasture)	Definition of unexposed group (access to pasture)	r/n in exposed group	r/n in unexposed group	Outcome	Measure of association	Covariates included	Measure of association value	95% confidence interval
(Gelasakis et al., 2013).	Zero grazing*	Grazing	NR/763	NR/356	Lameness (y/n)	Odds ratio	Parity, stocking density, lambing seasons, flock size	1.65	0.87-3.14 ²
	Zero grazing	Grazing	NR/763 ⁶	NR/356 ⁶	Severity of lameness	Odds ratio	As above	1.86 ¹	² 0.55-6.21
(Kaler and Green, 2009).	Housed for 1–2 months in 2003	Not housed in 2003	NR/205	NR/232	Lameness	Prevalence ratio	No covariates	0.94	0.80–1.11
	As above	As above	NR/111	NR/112	Interdigital dermatitis	Prevalence ratio	No covariates	1.12	0.85–1.46
	As above	As above	NR/111	NR/112	Footrot	Prevalence ratio	No covariates	0.99	0.72–1.35
(Kaler and Green, 2009).	Housed for >3 months in 2003	As above	NR/280	NR/232	Lameness	Prevalence ratio	No covariates	0.90	0.78–1.05
	As above	As above	NR/167	NR/112	Interdigital dermatitis	Prevalence ratio	No covariates	1.29	1.01–1.65
(Kaler and Green, 2009).	As above	As above	NR/167	NR/112	Footrot	Prevalence ratio	No covariates	1.34	1.01–1.78
(Wassink et al., 2004)	House ewes in winter ³	Do not house ewes in winter	NR/106	NR/41	Interdigital dermatitis in ewes	Odds ratio	No covariates	2.28	0.72–16.6
	As above	As above	NR/106	NR/40	Interdigital dermatitis in lambs	Odds ratio	No covariates	2.02	0.65–6.29
(Wassink et al., 2003)	Winter housing	No winter housing	NA	NA	Footrot	Rate ratio	No covariates	0.54 ⁴	0.33-0.86
(Green et al., 2004).	Housed for 3–6 months	Housed for <3 months	NR/NR	NR/NR	Footrot, Interdigital dermatitis	Odds ratio ⁵	No covariates	5.52	2.46–12.68

Citation	Exposed group (less or no access to pasture)	Definition of unexposed group (access to pasture)	r/n in exposed group	r/n in unexposed group	Outcome	Measure of association	Covariates included	Measure of association value	95% confidence interval
(Angell et al., 2014)	Housing not defined	Housing not defined	NR/NR	NR/NR	Contagious ovine digital dermatitis	NR	NR	NR	NR

r/n = Number of flocks (animals) with outcome/ the number in the group

NR= not reported

NA= not applicable – this study reported diseases rates

¹ Note the tables in this manuscript indicate that the referent is the access to pasture (grazing) and that access increases odds of occurrence by 1.65. However, other aspects of the paper suggest a transcription error and therefore the review authors have changed the referent here. Note that regardless of the referent this exposure was not associated with the outcome in this study.

² Indicates that these data are back calculated (based on assumptions) and therefore not reported by the authors

³ Highest levels of FR occurred during winter months (December to March)

⁴ Authors used “no winter housing group as a reference and an odds ratio of 1.85, however here it is translated to maintain consistent presentation with no pasture (or housed animals) as the referent.

⁵ Model description indicates a residual error with mean 0 and variance 1 for a logit transformed outcome for what appears to be a non-additive generalized linear model

⁶ these numbers are assume as the same groups were used

Table 14: Results from relevant studies with stocking density as an exposure identified in a systematic review of the effects of management system on lameness or lameness-related conditions in sheep in Europe.

Citation	Stocking density (higher)	Stocking density (lower)	r/n in exposed group	r/n in unexposed group	Outcome	Measure of association metric	Covariates	Measure of association value	Confidence interval
(Gelasakis et al., 2013).	≥2 m ² /ewe	<2 m ² /ewe	NR/593	NR/526	Lameness (y/n)	Odds ratio	Parity, stocking density, lambing seasons, flock size	2.31	1.23-4.14 ²
	≥2 m ² /ewe	<2 m ² /ewe	NR/NR	NR/NR	Severity of lameness	Odds ratio	As above	3.37 ¹	1.63-6.94 ²
(Kaler and Green, 2009).	>8 ewes/ha	≤8 ewes/ha	NR/101	NR/508	Lameness	Prevalence ratio	Frequency of foot trim , frequency of foot bathing, separate lame sheep at pasture, region,	1.26	1.05–1.50
	>8 ewes/ha	≤8 ewes/ha	NR/70	NR/302	Interdigital dermatitis	Prevalence ratio	As above	1.39	1.07–1.82
	>8 ewes/ha	≤8 ewes/ha	NR/70	NR/302	Footrot	Prevalence ratio	As above	0.94	0.70–1.22
(Wassink et al., 2004)	>8 ewes/ha	≤8 ewes/ha	NR/71	NR/73	Interdigital dermatitis in ewes	Odds ratio	No covariates	3.58	1.22–10.5
	>8 ewes/ha	≤8 ewes/ha	NR/71	NR/73	Interdigital dermatitis in lambs	Odds ratio	No covariates	1.88	0.65–5.43
(Wassink et al., 2003)	>8 ewes/ha	≤8 ewes/ha	NA	NA	Footrot	Rate ratio	No covariates	1.03	0.75–1.42

¹ Note the tables in this manuscript indicate that the referent is the higher stocking density and that lower density increase the odds of severity by 3.37. However, the discussion includes a long rationale for why higher stocking density would be associated with increased lameness. Based on the discussion and the arrangement of tables the review authors concluded that the tables contain a transcription error

² indicates that these data are back calculated (based on assumptions) and therefore not reported by the author

Table 15: Risk of bias assessment for observational studies included in the review

Risk of bias item	(Angell et al., 2014),	(Gelasakis et al., 2013)	(Green et al., 2004),	(Kaler and Green, 2009)	(Wassink et al., 2003)	(Wassink et al., 2004),
Do the inclusion/exclusion criteria vary across the comparison groups of the study?	No	Cannot determine:	Cannot determine:	No	No	No
Does the strategy for recruiting participants into the study differ across groups?	No	Cannot determine:	Cannot determine:	No	No	No
Is the selection of the comparison group inappropriate, after taking into account feasibility and ethical considerations?	No	No	Cannot determine	No	No	No
Does the study fail to account for important variations in the execution of the study from the proposed protocol?	NA	NA	NA	NA	NA	NA
Was the outcome assessor not blinded to the intervention or exposure status of participants?	Yes, not blinded	Yes, not blinded	Yes, not blinded	Yes, not blinded	Yes, not blinded	Yes, not blinded
Were valid and reliable measures, implemented consistently across all study participants used to assess inclusion/exclusion criteria, intervention/exposure outcomes, participant health benefits and harms, and confounding?	No	Cannot determine	Cannot determine	No	No	No
Was the length of follow-up different across study groups?	NA	No	Cannot determine	NA	NA	NA
In cases of high loss to follow-up was the impact assessed?	NA	No	Cannot determine	NA	NA	NA
Are any important primary outcomes missing from the results?	No	No	No	No	No	No
Are any important harms or adverse events that may be a consequence of the intervention/exposure missing from the results?	NA	NA	NA	NA	NA	NA
Was there any attempt to balance the allocation between the groups or match groups	Yes	Yes	Yes	Yes	Yes	Yes
Were important confounding variables not taken into account in the design and/or analysis?	Partially	Partially	Yes	No	Partially	No
Other biases: Control of non-independence	NA	Unclear if approach adjusted for non	Unclear if approach adjusted for non	NA	NA	NA

independe
nce

independen
ce

FIGURES

Figure 4: Flow diagram demonstrating the process of identification of relevant studies for synthesis

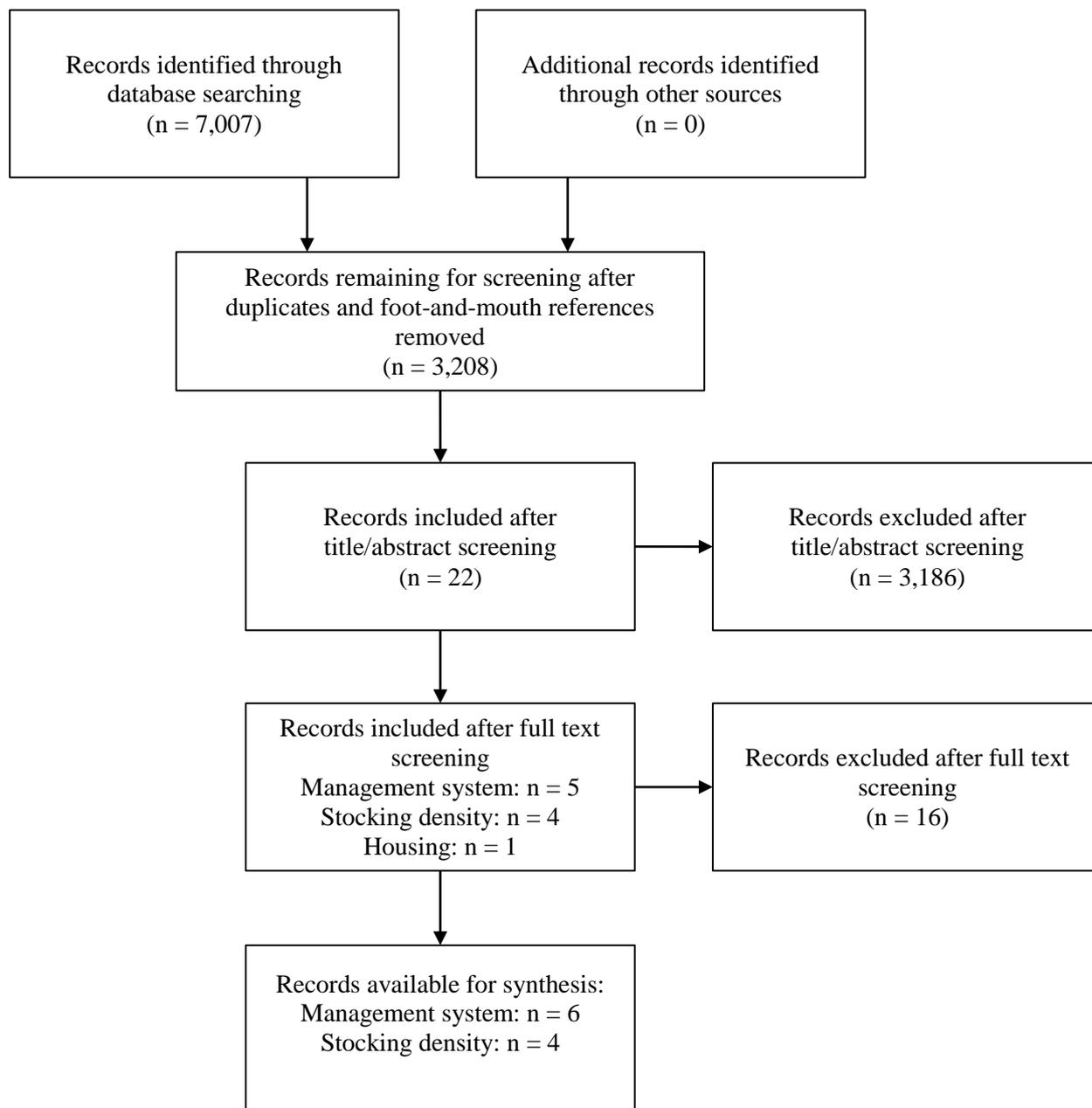


Figure 5: Search strategy to identify studies reporting on lameness in sheep in Science Citation Index (Web of Knowledge, Thompson Reuters)

```
# 10    #9 AND #1

# 9     #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2

# 8     TS=(polyarthriti* OR poly-arthriti* OR "stiff lamb disease" OR "erysipelas arthritis" OR "infectious
arthritis" OR "bacterial arthritis")

# 7     TS=(("septic" OR "sepsis") NEAR/3 ("pedal" OR "joint" OR "joints" OR "arthritis"))

# 6     TS=(("strawberry" OR "straw-berry" OR "shelly") NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR
"hoof" OR "hooves"))

# 5     TS=("digital dermatitis" OR "interdigital dermatitis" OR "CODD" OR "soil balling" OR "fibroma" OR
"fibromas" OR "granuloma" OR "granulomas" OR "laminitis" OR "founder" OR "joint ill")

# 4     TS=(footrot* OR foot-rot* OR "scald" OR "pododermatitis" OR "podo-dermatitis" OR "fusobacterium
necrophorum" OR "bacteroides melaninogenicus" OR "f necrophorum" OR "b melaninogenicus" OR
"dichelobacter nodosus" OR "d nodosus")

# 3     TS=((overgrow* OR over-grow* OR misshape* OR mis-shape* OR abscess* OR infect* OR disease*
OR injur* OR lesion* OR defect* OR inflam* OR "septic" OR "sepsis" OR "necrosis" OR "necroses" OR
degenerat*) NEAR/2 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves" OR "white line*" OR
"digital" OR "inter-digital"))

# 2     TS=("lame" OR "lameness")

# 1     TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR
"mutton" OR hogget* OR wether*)
```

Figure 6: Forest plot of association between different measures of lameness and variables that describe access to pasture. All data is organize such that the numerator of the association represents the animals housed or housed more frequently compared to the denominator which refers to animals with more access to pasture (less housing). A ratio greater than one suggests the numerator is a risk factor.

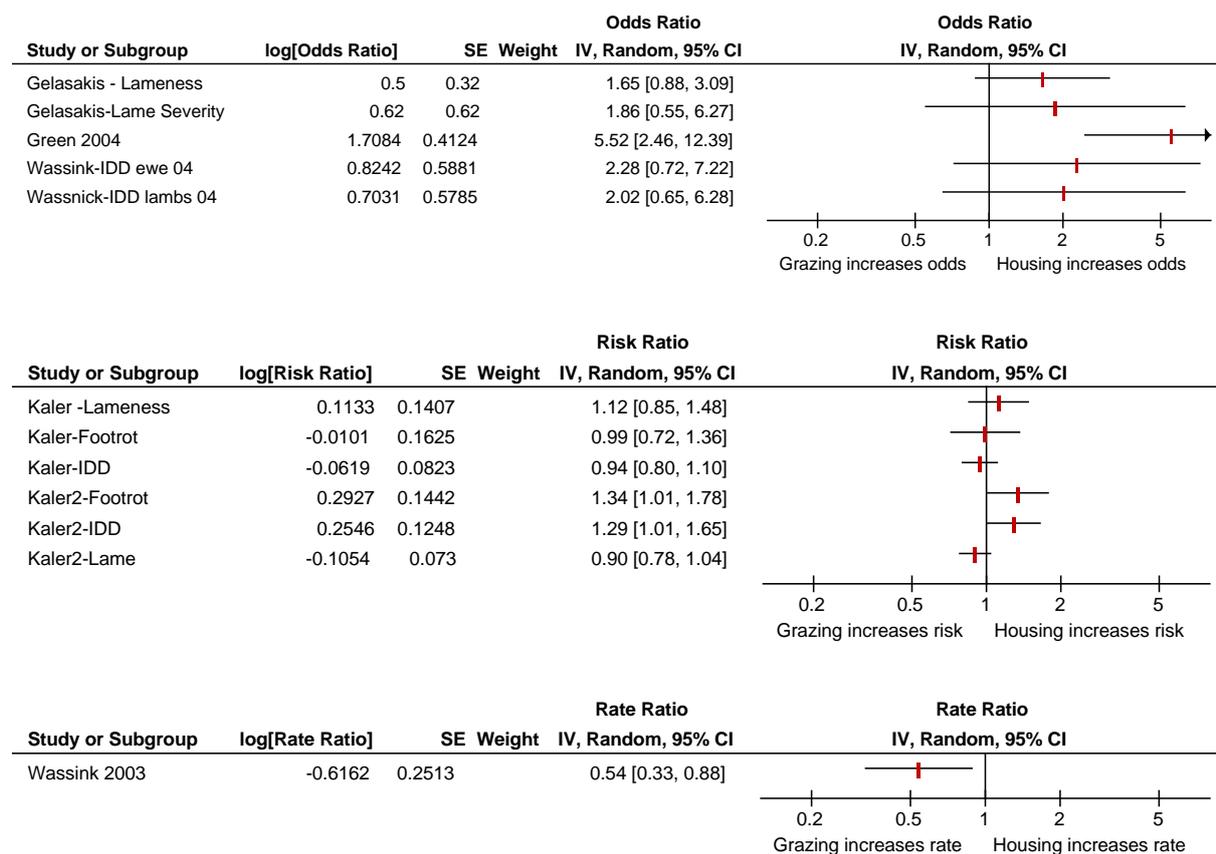


Figure 7: Forest plot of association between different measures of lameness and variables that describe access to pasture. The study by Green et al (Green et al., 2004) has been removed to provide illustration of the effect of that study on the interpretation of the data. The referent group is animals with no housing. A ratio greater than one suggests the numerator is a risk factor.

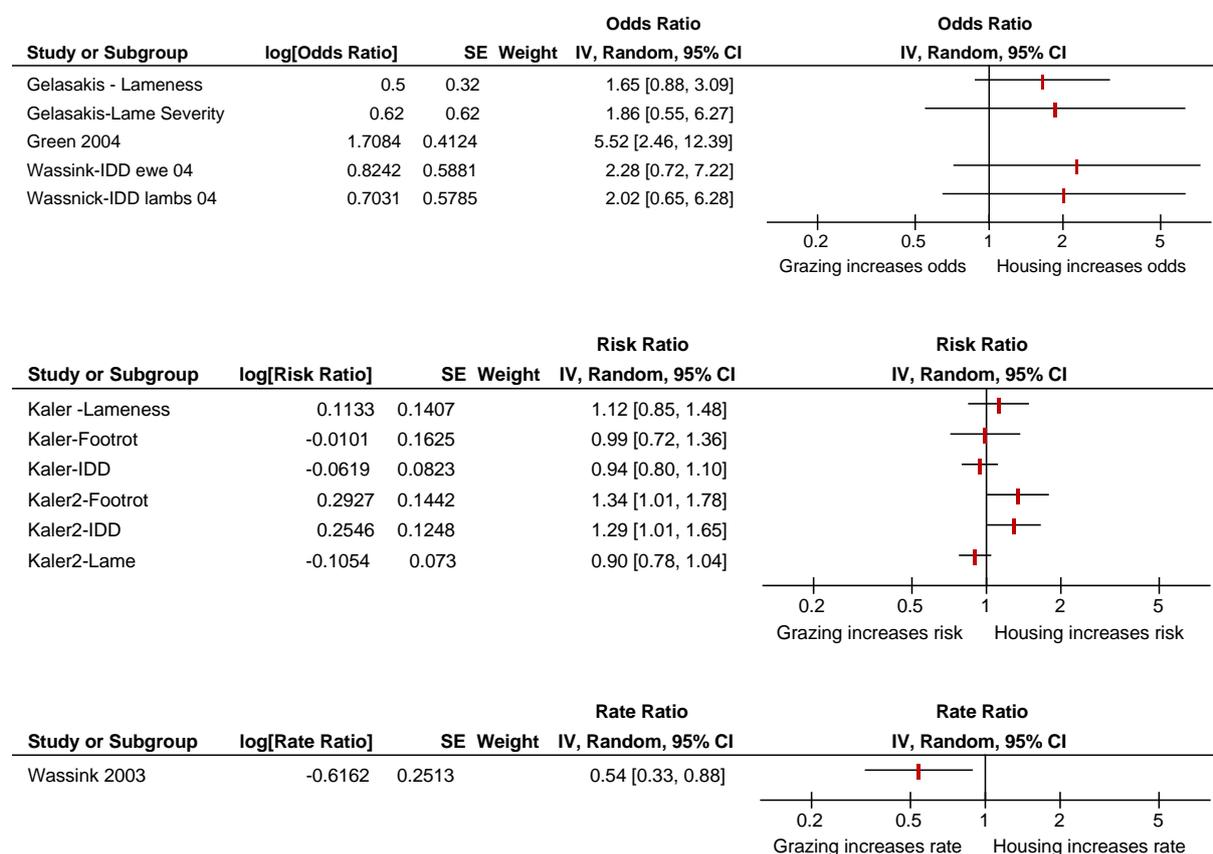
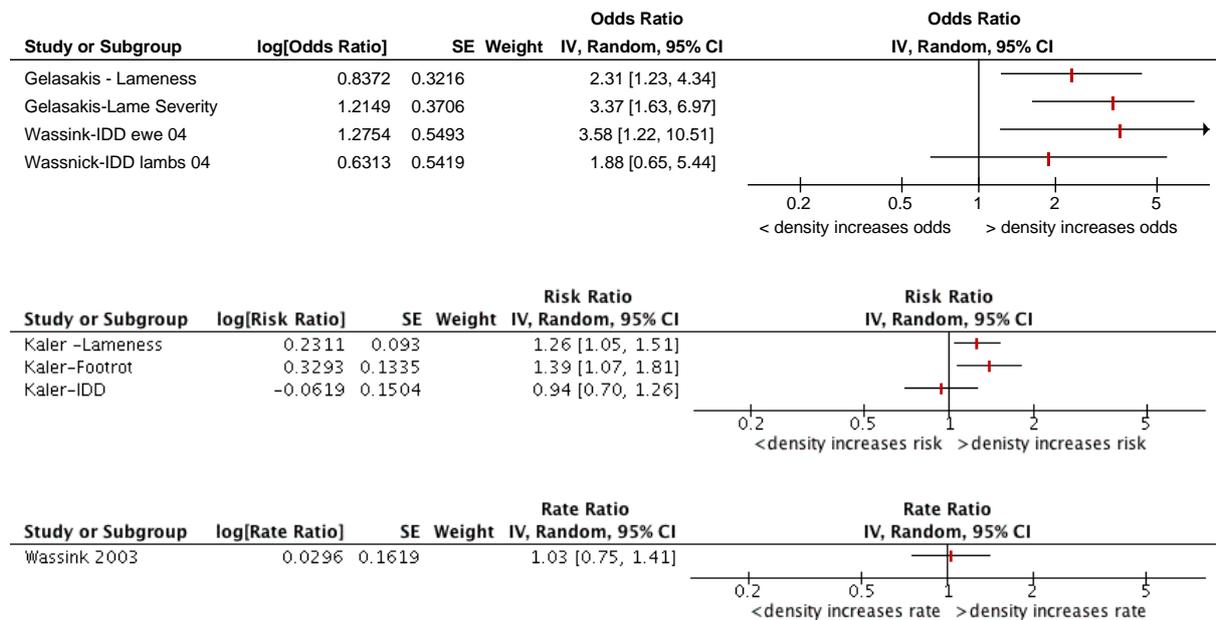


Figure 8: Forest plot of association between different measures of lameness and variables that describe stocking density. All data is organized such that the numerator of the association represents the high density compared to the denominator which refers to animals with lower density. A ratio greater than one suggests the numerator is a risk factor.



REFERENCES

- Angell, J.W., Duncan, J.S., Carter, S.D., Grove-White, D.H., 2014. Farmer reported prevalence and factors associated with contagious ovine digital dermatitis in Wales: A questionnaire of 511 sheep farmers. 113, 132-138.
- Gelasakis, A.I., Oikonomou, G., Bicalho, R.C., Valergakis, G.E., Fthenakis, G.C., Arsenos, G., 2013. Clinical characteristics of lameness and potential risk factors in intensive and semi-intensive dairy sheep flocks in Greece. *Journal of the Hellenic Veterinary Medical Society* 64, 123-130.
- Green, L.E., Wassink, G.J., Grogono-Thomas, R., Moore, L.J., Medley, G.F., 2004. Understanding the infectious aetiology of footrot in sheep: optimising interventions to minimise disease.
- Kaler, J., Green, L.E., 2009. Farmers' practices and factors associated with the prevalence of all lameness and lameness attributed to interdigital dermatitis and footrot in sheep flocks in England in 2004. 92, 52-59.
- RevMan, 2012. Review Manager (RevMan) [Computer program]. Version 5.2. The Cochrane Collaboration, 2012, Copenhagen: The Nordic Cochrane Centre.
- von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P., 2007. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 370, 1453-1457.
- Wassink, G.J., Grogono-Thomas, R., Moore, L.J., Green, L.E., 2003. Risk factors associated with the prevalence of footrot in sheep from 1999 to 2000. *Veterinary Record* 152, 351-358.
- Wassink, G.J., Grogono-Thomas, R., Moore, L.J., Green, L.E., 2004. Risk factors associated with the prevalence of interdigital dermatitis in sheep from 1999 to 2000. *Veterinary Record* 154, 551-+.

APPENDICES

APPENDIX FOR THE SCOPING REVIEW

Appendix A. Refinement of literature search for scoping review

Refinement of the search terms to be used for identification of the literature set for the scoping review—developing a full search strategy

An analysis of the clusters selected by the ISU and Guelph reviewers as containing potentially relevant citations showed that developing a search strategy with the appropriate balance of sensitivity and specificity in this topic area would particularly be problematic. Only a small percentage of potentially relevant citations explicitly made reference to welfare (or related terms such as quality of life, comfort, distress, pain etc.) in the title, abstract, or indexing of database records. The majority described alternative indicators for welfare in domestic sheep including

- physiological and metabolic responses such as temperature, respiration rate, and the release of biochemical stress indicators;
- behavioural indicators such as mother-lamb interactions, resting periods and head-shaking;
- rates of disease and injuries;
- physical condition including body weight and body condition.

To capture these studies, which do not explicitly discuss welfare, a very wide range of search terms were required. This included both generic terms (e.g. welfare, pain, stress, behaviour) and terms related to specific welfare indicators listed above (e.g. named sheep diseases or infections, unnatural behaviours such as hair pulling and head shaking, normal social interactions and behaviours; biochemical and physiological indicators such as cortisol levels, temperature, and heart rate).

Test searches developed using the most relevant clusters, in addition to the draft documentation provided by the EFSA Working Group on the Welfare of Sheep, comprised three key elements, using the Science Citation Index (Figure 4):

- sheep (search line 1);
AND
- welfare indicators (search lines 2 to 12);
AND
- terms associated with the production of sheep for wool, milk or meat. This includes terms related to feeding, housing, genetic factors, environmental factors, transportation and management procedures (search lines 13 to 19).

Figure 9: Initial sensitive search strategy constructed in Science Citation Index (Web of Knowledge, Thomson Reuters) 1900 to 22/01/14

```
# 21 #20 AND #12 AND #1 17,126
# 20 #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 4,197,752
# 19 TS=(("intensive" OR "intensively" OR "factory" OR "extensive" OR "extensively" OR "organic")
NEAR/2 ("system" OR "systems" OR "management" OR "farm" OR "farms" OR "farming" OR
"farmed")) 14,239
# 18 TS=("genetic selection" OR "selective breeding" OR "selectively bred" OR "genetic trait")
6,766
# 17 TS=(("feed*" OR "food" OR "water*" OR "drinking" OR "diet*" OR "nutrition*") NEAR/3
("space" OR "supply" OR "supplies" OR "access*" OR "availability" OR "scarcity" OR "restrict*" OR
"provision" OR "quality" OR "compete*" OR "competition" OR "deprive*" OR "deprivation"))
174,771
# 16 TS=("dog" OR "dogs" OR "predator" OR "predators" OR "predation" OR "social mixing" OR
"confinement" OR "flock size" OR "housing" OR "housed" OR "pen" OR "pens" OR "bunks" OR
"bedding" OR "litter" OR "space allowance*" OR "flooring" OR "shelter*" OR "shade" OR "air quality"
OR "ventilation" OR "ventilated" OR "ammonia concentrat*" OR "forage*" OR "foraging" OR "grazing"
OR "graze" OR "pasture" OR "pastures" OR "lighting" OR (("hot" OR "hotter" OR "cold" OR "colder")
NEAR/2 ("temperature*" OR "exposure*")) OR "climate" OR "climatic" OR "weather" OR "sun" OR
"snow" OR "blizzard*" OR "wind" OR "rain" OR "rainfall" OR "thermal" OR "thermoregulation" OR
"thermo-regulation" OR "humidity" OR "solar radiation" OR "drought" OR "exercise" OR "walking" OR
"environment" OR "environmental" OR "ecological") 3,079,585
# 15 TS=("transport" OR "transportation" OR "transit" OR "journey" OR "journeys" OR "long
distance*" OR "long haul" OR "crate" OR "crating" OR "over crowd*" OR "overcrowd*" OR "stock*
densit*" OR "trailer*" OR "lorry" OR "lorries" OR "truck*" OR "trailer" OR "ferry" OR "ferries" OR "boat"
OR "boats") 843,483
# 14 TS=("handling" OR "handler*" OR "restraint" OR "restraining" OR "restrained" OR (("human"
OR "humans" OR "staff" OR farmer* OR worker* OR "shepherd" OR "shepherds" OR "farmworker*"
OR "farm worker*" "farmhand*" OR "farm hand*") NEAR/4 (interact* OR contact* OR relationship* OR
exposure*)) 179,760
# 13 TS=("husbandry" OR "shepherding" OR "weaning" OR "gentling" OR "social isolation" OR
(("maternal*" OR "mother*") NEAR/3 ("deprivation" OR "isolation" OR "bonding" OR "separat*")) OR
"artificial* rear*" OR "artificial* feed*" OR "grooming" OR "shearing" OR "sheared" OR "shorn" OR
"driving" OR "droving" OR "rounding up" OR "crutching" OR "dagging" OR "mulesing" OR "mulsed"
OR "rubber ring" OR "castration" OR "castrating" OR "clipping" OR "clipped" OR "Burdizzo" OR
"docking" OR "docked" OR "ear tag*" OR "de-horn*" OR "dehorn*" OR "debudd*" or "disbudd*" OR
"tipping" OR "drench*" OR "dipping" OR "shower dip" OR "shower dips" OR "showers" OR
"showering") 235,138
# 12 #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 7,280,111
# 11 TS=("blood chemistry" OR "cortisol" OR "cortical" OR "corticosteroid*" OR
"adrenocorticotropic" OR "ACTH" OR "AVP" OR "corticotropin" OR "adrenocortical" OR "adrenal" OR
"adrenaline" OR "angiotensin" OR "aldosterone" OR "epinephrine" OR "norepinephrine" OR
"dopamine" OR "prolactin" OR "endorphin*" OR "glucocorticoid*" OR "heart rate*" OR "heartrate*" OR
"respiration rate*" OR "respiratory rate*" OR (("ear" OR "ears" OR "rectal*" OR "rectum") NEAR/3
"temperature*"))865,829
# 10 TS=("wasted" OR "wasting" OR "cachexia" OR "anorexia" OR "starving" OR "starvation" OR
"starved" OR "body condition" OR "muscle condition" OR "muscle atrophy" OR "malnutrition" OR
"undernutrition" OR "hunger" OR "hungry" OR "thirst*" OR "dehydrat*" OR "deficien*") 627,290
# 9 TS=("scrapie" OR "milk fever" OR "paresis" OR "hypocalcaemia" OR "hypocalcemia" OR
"stagger*" OR "grass tetany" OR "winter tetany" OR "hypomagnes*" OR "sleepy sickness" OR "twin
lamb disease" OR "aceto-naemia" OR "acetonemia" OR "coccidiosis" OR "coccidia" OR
"toxoplasmosis" OR "campylobacter*" OR "vibriosis" OR "listeria" OR "listeriosis" OR "circling" OR
"polio*" OR "clostridium" OR "clostridial" OR "enteritis" OR "scour" OR "scours" OR "diarrhea" or
"diarrhoea" OR "tetanus" OR "lockjaw" OR "lock jaw" OR "anthrax" OR "visna virus" OR "visna-maedi")
```

OR "maedi-visna" OR "ovine lentivirus" OR "salmonell*" OR "chlamydia" OR "acidosis" OR "acidois")
355,726

8 TS=("broken mouth*" OR "sore mouth*" OR "soremouth*" OR "ecthyma" OR "footrot*" OR "foot rot" OR "overgrown hoof" OR "overgrown hooves" OR "pododermatitis" OR "scald*" OR "dermatitis" OR "pink eye*" OR "pinkeye*" OR "keratoconjunctivitis" OR "conjunctivitis" OR "blepharoconjunctivitis" OR "mastitis" OR "hard bag" OR "blue bag" OR "agalactia" OR "agalactorrhea" OR "agalactorrhoea" OR "flystrike*" OR "fly strike*" OR "flystruck" OR "fly struck" OR "flyblown" OR "fly blown" OR "blowfly strike*" OR "blowfly struck" OR "myiasis" OR "lucilia sericata" OR "pneumonia" OR "bronchopneumonia" OR "pasteurellosis" OR "shipping fever*" OR "paratuberculosis" OR "para-tuberculosis" OR "Johne's" OR "Johnes" OR "scab" OR "mite" OR "mites" OR "psoroptes" OR "pox" OR "parasit*" OR "ectoparasit*" OR "worm*" OR "tapeworm*" OR "lungworm*" OR "hydatids" OR "louse" OR "lice" OR "bovicola ovis" OR "sheep ked*" OR "tick" OR "ticks" OR "melophagus ovinus" OR "bot" OR "bots" OR "mange") 434,533

7 TS=("abortion*" OR "stillbirth*" OR "still birth*" OR "stillborn" OR "still born" OR ("obstructed" OR "prolong*") NEAR/3 ("labor*" OR "labour*" OR "birth*")) OR "dystocia" OR "mismother*" OR "mismother*" OR "deserted" OR "deserting" OR "bonding") 200,755

6 TS=("disease" OR "diseases" OR "poison*" OR "toxin*" OR "toxemia" OR "toxaemia" OR "endotoxin*" OR "endotoxemia" OR "endotoxaemia" OR "infection*" OR "infectious" OR "abscess*" OR "necrosis" OR "immune system" OR "immune response*" OR "injury" OR "injured" OR "injuries" OR "wound*" OR "fracture*" OR "lame" OR "lameness" OR "smother*" OR "crushing" OR "crushed" OR "hyperthermi*" OR "hypothermi*") 2,475,114

5 TS(("metabolic" OR "metabolism" OR "psychological" OR "biochemical" OR "physiological" OR "physical" OR "hormon*") NEAR/1 ("effect" OR "effects" OR "influence*" OR "response*" OR "consequence*" OR "impact" OR "adapt*" OR "marker*")) 141,511

4 TS(("hair" OR "wool") NEAR/3 ("eat" OR "eats" OR "eating" OR "eaten" OR "pull" OR "pulled" OR "pulls" OR "pulling" OR "bite" OR "biting" OR "bitten")) OR (("bar" OR "bars" OR "pen" OR "pens" OR "wall" OR "walls") NEAR/3 ("chew*" OR "bite" OR "biting" OR "bitten" OR "lick*" OR "gnaw*")) OR "stargaz*" OR "star-gaz" OR "pacing" OR "route tracing" OR ("head" OR "heads") NEAR/3 (shake* OR shaking*)) 25,490

3 TS=("behavior*" OR "behaviour*" OR "social interaction*" OR "social contact*" OR "visual contact*" OR "vocal contact*" OR "vocalis*" OR "vocaliz*" OR "bleat*" OR "aggress*" OR "fight*" OR "bully*" OR "resting" OR "frustrat*" OR "agitat*" OR "aversion") 2,017,124

2 TS=("welfare" OR "stress" OR "stresses" OR "stressor*" OR "distress*" OR "pain" OR "painful" OR "fear*" OR "comfort" OR "discomfort" OR "depression" OR "depressed" OR "temperament" OR "temperaments" OR "emotion" OR "emotional") 1,666,894

1 TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR "hogget*") 169,821

An initial draft of this approach returned approximately 17,000 records, a volume beyond that which can be managed within the constraints of this project. Many of these results were non-relevant, retrieved by the wide variety of terms necessary to capture the second and third key elements. Many of the most frequently used terms identified by the clustering activity such as pain, behaviour, stress, farming, husbandry, genetics, and environment are non-specific and return many non-relevant records even when combined with the other key elements of the search.

The problem was compounded by the somewhat limited search functionality of the key information resources relevant to this topic area: Science Citation Index and CAB Abstracts via Web of Knowledge. The Web of Knowledge interface does not provide the option to restrict a search to the title and abstract fields; to search abstract fields, you must search by "topic" and this includes many superfluous fields that increase the volume of non-relevant results returned. This issue would be even more apparent in other agriculture or animal health databases such as AGRICOLA and AGRIS which do not allow the type of complex, multi-stranded searches required by the likely review question(s).

Appendix B. The abstract screening and abstract characterization questions for scoping review

Question Text	Answer Text
Does this primary study describe a potential determinant of welfare and a change in a welfare outcome of sheep used in wool, meat, or milk production on farm?	No
	Yes
	Yes, but no links reported between potential determinant(s) of welfare and outcome(s)
	Potentially relevant review, book/chapter, or report of an on-farm practice or on-farm welfare issue
	Potentially relevant conference proceeding
	Potentially relevant, no abstract or abstract lacking detail
Is the full paper available in English?	Yes
	No
	Can't tell
Which of the following welfare determinants was/were studied or reported as being potentially associated with a change in a welfare outcome of sheep on farm?	Genetics
	Nutrition/feeding/watering, including all issues related to intake (i.e., food type, food/water volume intake, trough types, water source types, etc.)
	Housing: all STRUCTURAL components of the indoor environment that are not related to food or water or ventilation (e.g., flooring, pen construction, bedding)
	Environment, including all NON-STRUCTURAL aspects of indoor environments that are not related to food or water (e.g., ventilation, temperature, sound, lighting, stocking density) and all aspects of outdoor environments not related to food or water (e.g., terrain, temperature, shade, stocking density). Environment also includes those management practices related to those elements described.
	Management practices not related to those described under Nutrition/feeding/watering or Housing or Environment (e.g., comparison of methods for castration, weaning, shearing; isolation vs. group confinement; etc.)
	Health: specific diseases that potentially influence welfare (e.g., parasitism, dystocia)
	Behaviour (e.g., mis-mothering, aggression/bullying)
	Handler traits (e.g., training, competence, gender, lack of sympathy with sheep)
What welfare outcome(s) was/were evaluated?	Abnormal clinical sign/physiological measurement, without a specified disease (e.g., tachycardia, dehydration)
	Abnormal haematology or blood biochemistry

	measurement, without without a specified disease (e.g., elevated cortisol, decreased lymphocytes)
	Impaired maternal/mothering/lamb behaviours (includes lamb stealing, increased time to stand/suckle, and subjective lamb vigour scores and maternal behaviour scores)
	Impaired resting behaviours
	Abnormal social behaviour (e.g., aggression, agonistic/bullying)
	Other abnormal behavior includes vocalization
	Death
	Dystocia
	Lameness
	Fly strike
	Impaired/delayed wound healing
	Other dermatological disease/injury, including exoparasites and footrot
	Other musculoskeletal disease/injury
	Other reproductive disease/injury (includes duration of parturition)
	Poor body condition
	Birth defects/congenital abnormalities
	Cardiac disease
	Gastrointestinal disease/injury, including GIT parasites
	Immunological disease
	Mammary disease/injury
	Neurological disease
	Ocular disease/injury
	Oral disease/injury
	Respiratory disease
	Septicemia
	Urinary disease
	Other disease/injury
	Distress (the authors explicitly use the word "distress")
	Suffering (the authors explicitly use the word "suffering")
	Stress (behavioural, physiological, or biochemical measurement, or just said they evaluated stress)
	Fear (behavioural, physiological, or biochemical measurement, or just said they evaluated fear or aversion)
	Pain (behavioural, physiological, or biochemical measurement, or just said they evaluated pain)
	Thermal stress (behavioural, physiological, or biochemical measurement, or just said they evaluated thermal stress)
What population was studied?	Undifferentiated flock(s)
	Ewes
	Lambs
	Ewes and lambs
	Rams

	Rams, ewes, and lambs
	Wethers
	Adults
	Not discernible
What study design was used?	Observational/descriptive study (either no control group or no imposed intervention)
	Controlled trial (imposed intervention and at least 2 comparative groups)
	Not discernible
What production system(s) was/were studied?	Milk
	Meat
	Wool
	Not described
What farm management system was studied?	Extensive (outdoor observational studies and experimental studies explicitly evaluating a component of the outdoor environment (e.g., the effect of shelters on the survival of lambs born outside))
	Intensive (indoor observational studies and experimental studies explicitly evaluating a component of indoor housing (e.g., flooring))
	Mixed
	Not explicitly described - Observational studies
	Not applicable, experimental study
In what country did the study occur?	Not Reported
	Reported
Comments:	

Appendix C. Table describing the number of studies in each welfare determinate and the frequency of outcomes reported in that determinate for scoping review.

Studies may appear in more than one determinate.

Genetics (146 studies)	Frequency of outcome
Abnormal clinical sign	45
Abnormal blood/urine/etc. indicator	37
Maternal/lamb behaviour	62
Resting behaviour	10
Abnormal social behaviour	10
Abnormal behaviour	34
Death	47
Dystocia	11
Lameness	1
Fly strike	1
Non-fly-strike dermatological	1
Non-dystocia reproductive	13
Congenital abnormality	1
Gastrointestinal/GI parasites	2
Mammary	1
Neurological	1
Respiratory	1
Other disease/injury	2
Distress	3
Stress	4
Fear	4
Pain	1
Thermal stress	15
Behaviour (56 studies)	
Abnormal clinical sign	8
Abnormal blood/urine/etc. indicator	11
Maternal/lamb behaviour	36
Resting behaviour	3
Abnormal social behaviour	7
Abnormal behaviour	14
Death	26
Dystocia	1
Non-dystocia reproductive	2

Congenital abnormality	1
Mammary	1
Stress	3
Fear	3
Pain	2
Thermal stress	3
Environment (215 studies)	
Abnormal clinical sign	124
Abnormal blood/urine/etc. indicator	98
Maternal/lamb behaviour	12
Resting behaviour	9
Abnormal social behaviour	6
Abnormal behaviour	36
Death	30
Dystocia	4
Non-fly-strike dermatological	1
Non-lameness musculoskeletal	3
Non-dystocia reproductive	5
Congenital abnormality	3
Gastrointestinal/GI parasites	3
Neurological	1
Respiratory	5
Other disease/injury	5
Distress	1
Stress	7
Fear	4
Thermal stress	32
Handler (15 studies)	
Abnormal clinical sign	4
Abnormal blood/urine/etc. indicator	5
Resting behaviour	2
Abnormal social behaviour	2
Abnormal behaviour	13
Distress	1
Stress	1
Fear	4
Pain	1
Housing (39 studies)	
Abnormal clinical sign	10
Abnormal blood/urine/etc. indicator	13
Maternal/lamb behaviour	5

Resting behaviour	6
Abnormal social behaviour	6
Abnormal behaviour	12
Death	7
Dystocia	1
Non-fly-strike dermatological	1
Non-lameness musculoskeletal	2
Non-dystocia reproductive	1
Respiratory	2
Other disease/injury	1
Stress	3
Thermal stress	2
Management (343 studies)	
Abnormal clinical sign	101
Abnormal blood/urine/etc. indicator	193
Maternal/lamb behaviour	40
Resting behaviour	44
Abnormal social behaviour	28
Abnormal behaviour	145
Death	20
Dystocia	1
Lameness	3
Fly strike	4
Impaired/delayed wound healing	11
Non-fly-strike dermatological	5
Non-lameness musculoskeletal	1
Non-dystocia reproductive	2
Congenital abnormality	1
Gastrointestinal/GI parasites	7
Immunological	1
Mammary	3
Neurological	1
Respiratory	2
Other disease/injury	5
Distress	17
Stress	32
Fear	13
Pain	33
Thermal stress	11
Nutrition	
Abnormal clinical sign	34

Abnormal blood/urine/etc. indicator	46
Maternal/lamb behaviour	19
Resting behaviour	6
Abnormal social behaviour	5
Abnormal behaviour	28
Death	15
Dystocia	2
Non-fly-strike dermatological	1
Non-lameness musculoskeletal	1
Non-dystocia reproductive	5
Gastrointestinal/GI parasites	1
Neurological	2
Oral	1
Respiratory	1
Other disease/injury	3
Stress	6
Fear	1
Thermal stress	8

APPENDIX FOR THE SYSTEMATIC REVIEW

Appendix D. Search strategies for systematic review

A1. Social Science Citation Index – Expanded (Web of Science, Thomson Reuters) 1900-2014-04-02. Searched 03/04/14

10 1,377 #9 AND #1

9 76,557 #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2

8 6,063 TS=(polyarthriti* OR poly-arthritis* OR "stiff lamb disease" OR "erysipelas arthritis" OR "infectious arthritis" OR "bacterial arthritis")

7 5,012 TS=("septic" OR "sepsis") NEAR/3 ("pedal" OR "joint" OR "joints" OR "arthritis")

6 15 TS=("strawberry" OR "straw-berry" OR "shelly") NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves")

5 41,666 TS=("digital dermatitis" OR "interdigital dermatitis" OR "CODD" OR "soil balling" OR "fibroma" OR "fibromas" OR "granuloma" OR "granulomas" OR "laminitis" OR "founder" OR "joint ill")

4 4,339 TS=(footrot* OR foot-rot* OR "scald" OR "pododermatitis" OR "podo-dermatitis" OR "fusobacterium necrophorum" OR "bacteroides melaninogenicus" OR "f necrophorum" OR "b melaninogenicus" OR "dichelobacter nodosus" OR "d nodosus")

3 14,550 TS=((overgrow* OR over-grow* OR misshape* OR mis-shape* OR abscess* OR infect* OR disease* OR injur* OR lesion* OR defect* OR inflam* OR "septic" OR "sepsis" OR "necrosis" OR "necroses" OR degenerat*) NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves" OR "white line*" OR "digital" OR "inter-digital"))

2 7,274 TS=("lame" OR "lameness")

1 172,117 TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR hogget* OR wether*)

A2. Conference Proceedings Citation Index – Science (Web of Science, Thomson Reuters) 1900-2014-04-02. Searched 03/04/14

10 68 #9 AND #1

9 4,813 #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2

8 280 TS=(polyarthriti* OR poly-arthritis* OR "stiff lamb disease" OR "erysipelas arthritis" OR "infectious arthritis" OR "bacterial arthritis")

7 242 TS=("septic" OR "sepsis") NEAR/3 ("pedal" OR "joint" OR "joints" OR "arthritis")

6 0 TS=("strawberry" OR "straw-berry" OR "shelly") NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves")

5 2,151 TS=("digital dermatitis" OR "interdigital dermatitis" OR "CODD" OR "soil balling" OR "fibroma" OR "fibromas" OR "granuloma" OR "granulomas" OR "laminitis" OR "founder" OR "joint ill")

4 412 TS=(footrot* OR foot-rot* OR "scald" OR "pododermatitis" OR "podo-dermatitis" OR "fusobacterium necrophorum" OR "bacteroides melaninogenicus" OR "f necrophorum" OR "b melaninogenicus" OR "dichelobacter nodosus" OR "d nodosus")

3 1,189 TS=((overgrow* OR over-grow* OR misshape* OR mis-shape* OR abscess* OR infect* OR disease* OR injur* OR lesion* OR defect* OR inflam* OR "septic" OR "sepsis" OR "necrosis" OR "necroses" OR degenerat*) NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves" OR "white line*" OR "digital" OR "inter-digital"))

2 639 TS=("lame" OR "lameness")

1 18,574 TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR hogget* OR wether*)

A3. Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present> Searched 03/04/14

1 exp Sheep/ (102629)

2 Sheep Diseases/ (18324)

3 (sheep or ovine or lamb or lambs or ewe or ewes or ram or rams or mutton or hogget\$ or wether\$).ti,ab,kf. (105924)

4 or/1-3 (134928)

5 Foot Diseases/ (10447)

6 Lameness, Animal/ (3233)

7 "Hoof and Claw"/ab, in, pa [Abnormalities, Injuries, Pathology] (646)

8 ((overgrow\$ or over-grow\$ or misshape\$ or mis-shape\$ or abscess\$ or infect\$ or disease\$ or injur\$ or lesion\$ or defect\$ or inflam\$ or septic or sepsis or necrosis or necroses or degenerat\$) adj3 (foot or feet or toe or toes or hoof or hooves or white line\$ or digital or inter-digital)).ti,ab,kf. (14731)

9 (footrot\$ or foot-rot\$ or scald or pododermatitis or podo-dermatitis or fusobacterium necrophorum or bacteroides melaninogenicus or f necrophorum or b melaninogenicus or dichelobacter nodosus or d nodosus).ti,ab,kf. (3448)

10 (digital dermatitis or interdigital dermatitis or CODD or soil balling or fibroma or fibromas or granuloma or granulomas or laminitis or founder or joint ill).ti,ab,kf. (46910)

11 ((strawberry or straw-berry or shelly) adj3 (foot or feet or toe or toes or hoof or hooves)).ti,ab,kf. (7)

12 ((septic or sepsis) adj3 (pedal or joint or joints or arthritis)).ti,ab,kf. (4648)

- 13 (polyarthriti\$ or poly-arthriti\$ or stiff lamb disease or erysipelas arthritis or infectious arthritis or bacterial arthritis).ti,ab,kf. (8502)
- 14 Foot Rot/ (362)
- 15 Digital Dermatitis/ or Foot Dermatoses/ (4267)
- 16 exp Arthritis, Infectious/ve [Veterinary] (588)
- 17 or/5-16 (92107)
- 18 4 and 17 (1371)
- 19 exp africa/ or exp americas/ or antarctic regions/ or arctic regions/ or exp asia/ or exp australia/ or exp oceania/ (2138111)
- 20 exp Europe/ (1089482)
- 21 exp asia, central/ or exp asia, northern/ or israel/ or turkey/ (86971)
- 22 19 not (20 or 21) (1968845)
- 23 18 not 22 (1202)

A4. Biosis Citation Index (Web of Science – Thomson Reuters) 1969-current. Searched 03/04/14

10

1,500

#9 AND #1

9

67,397

#8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2

8

5,078

TS=(polyarthriti* OR poly-arthriti* OR "stiff lamb disease" OR "erysipelas arthritis" OR "infectious arthritis" OR "bacterial arthritis")

7

3,056

TS=(("septic" OR "sepsis") NEAR/3 ("pedal" OR "joint" OR "joints" OR "arthritis"))

6

7

TS=("strawberry" OR "straw-berry" OR "shelly") NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves")

5

35,086

TS=("digital dermatitis" OR "interdigital dermatitis" OR "CODD" OR "soil balling" OR "fibroma" OR "fibromas" OR "granuloma" OR "granulomas" OR "laminitis" OR "founder" OR "joint ill")

4

5,217

TS=(footrot* OR foot-rot* OR "scald" OR "pododermatitis" OR "podo-dermatitis" OR "fusobacterium necrophorum" OR "bacteroides melaninogenicus" OR "f necrophorum" OR "b melaninogenicus" OR "dichelobacter nodosus" OR "d nodosus")

3

16,887

TS=((overgrow* OR over-grow* OR misshape* OR mis-shape* OR abscess* OR infect* OR disease* OR injur* OR lesion* OR defect* OR inflam* OR "septic" OR "sepsis" OR "necrosis" OR "necroses" OR degenerat*) NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves" OR "white line*" OR "digital" OR "inter-digital"))

2

3,967

TS=("lame" OR "lameness")

1

182,249

TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR hogget* OR wether*)

A5. CAB Abstracts (Web of Science – Thomson Reuters) 1910-current. Searched 03/04/14

Note: Due to the large volume of records identified in CAB Abstracts the strategy was refined to include a third concept; geographic setting. Studies with a non-European country in the descriptor field were excluded.

12

2,019

#10 not #11

11

2,185,724

DE=(Australia OR New Zealand OR United States OR USA OR America OR Brazil OR Egypt OR India OR Africa OR Algeria OR Argentina OR Cameroon OR Chile OR China OR Ghana OR Nigeria OR Saudi Arabia OR Sudan OR Canada OR Japan)

10

2,898

#9 AND #1

9

43,110

#8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2

8

1,529

TS=(polyarthriti* OR poly-arthritis* OR "stiff lamb disease" OR "erysipelas arthritis" OR "infectious arthritis" OR "bacterial arthritis")

7

996

TS=("septic" OR "sepsis") NEAR/3 ("pedal" OR "joint" OR "joints" OR "arthritis")

6

33

TS=("strawberry" OR "straw-berry" OR "shelly") NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves")

5

16,857

TS=("digital dermatitis" OR "interdigital dermatitis" OR "CODD" OR "soil balling" OR "fibroma" OR "fibromas" OR "granuloma" OR "granulomas" OR "laminitis" OR "founder" OR "joint ill")

4

10,071

TS=(footrot* OR foot-rot* OR "scald" OR "pododermatitis" OR "podo-dermatitis" OR "fusobacterium necrophorum" OR "bacteroides melaninogenicus" OR "f necrophorum" OR "b melaninogenicus" OR "dichelobacter nodosus" OR "d nodosus")

3

4,263

TS=((overgrow* OR over-grow* OR misshape* OR mis-shape* OR abscess* OR infect* OR injur* OR lesion* OR defect* OR "septic" OR "sepsis" OR "necrosis" OR "necroses" NEAR/3 ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves" OR "white line*" OR "digital" OR "inter-digital"))

2

12,130

TS=("lame" OR "lameness")

1

236,879

TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR hogget* OR wether*)

A6. International Information System for the Agricultural Sciences and Technology [AGRIS] 1975 to date <http://agris.fao.org/> Searched 07/04/14

Can only download 250 records at one time – search broken down into smaller segments to accommodate this.

(sheep ovine lamb lambs ewe ewes ram rams mutton hogget* wether*) AND (lame lameness) 96 results

(sheep ovine lamb lambs ewe ewes ram rams mutton hogget* wether*) AND (pododermatitis pododermatitis "digital dermatitis" "interdigital dermatitis" CODD) 33 results

(sheep ovine lamb lambs ewe ewes ram rams mutton hogget* wether*) AND (overgrow* over-grow* misshape* mis-shape* abscess* lesion*) AND (foot feet toe toes hoof hooves "white line*" digital inter-digital) 109 results

(sheep ovine lamb lambs ewe ewes ram rams mutton hogget* wether*) AND (shelly strawberry scald laminitis founder) 52 results

(sheep ovine lamb lambs ewe ewes ram rams mutton hogget* wether*) AND ("erysipelas arthritis" OR "infectious arthritis" OR "bacterial arthritis" OR polyarthritis OR "septic pedal") 60 results

(sheep ovine lamb lambs ewe ewes ram rams mutton hogget* wether*) AND (fibroma fibromas granuloma granulomas) 49 results

(sheep ovine lamb lambs ewe ewes ram rams mutton hogget* wether*) AND (footrot "foot rot") AND (incidence prevalence epidemiology epidemiological observational morbidity mortality burden) 67 results -- foot rot has to be limited otherwise thousands of records that cannot be exported.

**A7. National Agriculture Library Catalog [AGRICOLA] 1970-Current
<http://agricola.nal.usda.gov/> Searched 07/04/14**

Search Articles Database: Advanced Search

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR hogget? OR wether?)AND(lame OR lameness) 82 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget? OR wether?) AND (pododermatitis OR podo-dermatitis OR "digital dermatitis" OR "interdigital dermatitis" OR CODD) 29 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget? OR wether?)AND(overgrow? OR misshape? OR abscess? OR lesion? OR "over grow?" OR "mis shape?") AND (foot OR feet OR toe OR toes OR hoof OR hooves OR "white line?" OR digital OR "inter digital") 82 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget? OR wether?)AND(shelly OR strawberry OR scald OR laminitis OR founder) 47 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget? OR wether?)AND(fibroma? OR granuloma?)AND (feet OR foot OR toe OR toes OR hoof OR hooves) 4 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget? OR wether?)AND("erysipelas arthritis"OR"infectious arthritis"OR"bacterial arthritis"OR polyarthritis OR "septic pedal") 29 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget? OR wether?)AND(footrot OR "foot rot")AND(incidence OR prevalence OR epidemiolog? OR observational OR morbidity OR mortality OR burden) 46 results

**A8. TEKTRAN: The ARS Manuscripts Database <http://www.ars.usda.gov/services/tektran.htm>
Searched 11/04/13 Searched 08/04/14**

Browse: Measure & Evaluate Animal Well-Being, Animal Disease Epidemiology

Search: lame, lameness

Records manually scanned; 0 records added to EndNote

**A9. National Institute of Food and Agriculture Current Research Information System [CRIS]
<http://cris.nifa.usda.gov/> Searched 08/04/14**

CRIS Assisted Search (automatic truncation)

Fulltext Terms: lame, lameness AND sheep, ovine, lamb, ewe, ram, mutton, hogget, wether

Records manually scanned; 0 records added to EndNote

A10. Open Grey <http://www.opengrey.eu/> Searched 08/04/13

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR hogget* OR wether*) AND (lame OR lameness) 4results – 2 obviously irrelevant, 2 added to EndNote.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (pododermatitis OR podo-dermatitis OR "digital dermatitis" OR "interdigital dermatitis" OR CODD) 0 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (overgrow* OR misshape* OR abscess* OR lesion* OR "over grow*" OR "mis shape*") AND (foot OR feet OR toe OR toes OR hoof OR hooves OR "white line*" OR digital OR "inter digital") 1 result, obviously irrelevant, not downloaded.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (shelly OR strawberry OR scald OR laminitis OR founder) 2 results, obviously irrelevant, not downloaded.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (fibroma* OR granuloma*) 1 result, obviously irrelevant, not downloaded.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (arthritis OR polyarthritis OR "septic pedal") 2 results, 1 obviously irrelevant, 1 added to EndNote

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (footrot OR "foot rot") 2 results, 1 obviously irrelevant, 1 duplicate, 0 records added to EndNote.

A11. Science.gov <http://www.science.gov/> Searched 08/04/13

Advanced search: Full Record:

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR hogget* OR wether*) AND (lame OR lameness) (In: Agriculture & Food, Biology & Nature – NOT TEKTRAN and AGRICOLA which have already been search via their native interfaces) 2 results, 2 obviously irrelevant, 0 downloaded to EndNote.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (pododermatitis OR podo-dermatitis OR "digital dermatitis" OR "interdigital dermatitis" OR CODD) (In: Agriculture & Food, Biology & Nature – NOT TEKTRAN and AGRICOLA which have already been search via their native interfaces) 2 results, 2 obviously irrelevant, 0 downloaded to EndNote.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (overgrow* OR misshape* OR abscess* OR lesion* OR "over grow*" OR "mis shape*") AND (foot OR feet OR toe OR toes OR hoof OR hooves OR "white line*" OR digital OR

"inter digital") (In: Agriculture & Food, Biology & Nature – NOT TEKTRAN and AGRICOLA which have already been search via their native interfaces) 0 results

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (shelly OR strawberry OR scald OR laminitis OR founder) (In: Agriculture & Food, Biology & Nature – NOT TEKTRAN and AGRICOLA which have already been search via their native interfaces) 29 results – all irrelevant records related to strawberries, 0 added to EndNote.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (fibroma* OR granuloma*) (In: Agriculture & Food, Biology & Nature – NOT TEKTRAN and AGRICOLA which have already been search via their native interfaces) 1 result, obviously irrelevant, 0 records added to EndNote.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (arthritis OR polyarthritis OR "septic pedal") (In: Agriculture & Food, Biology & Nature – NOT TEKTRAN and AGRICOLA which have already been search via their native interfaces) 18 results, all obviously irrelevant, 0 records added to EndNote.

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR mutton OR hogget* OR wether*) AND (footrot OR "foot rot") 0 results.

A12. ScienceResearch.com <http://www.scienceresearch.com/> Searched 08/04/14

Advanced search: Full Record:

(sheep OR ovine OR lamb OR lambs OR ewe OR ewes OR ram OR rams OR hogget* OR wether*) AND (lame OR lameness) (Agriculture Sciences, Biology & Nature)

Topic cluster: Lameness in sheep: 61 results

Topic cluster: Ovine footrot: 6 results

Topic cluster: Limb lameness: 6 results

Appendix E. Study Selection questions for systematic review

Question 1: Does the title/abstract describe an observational study evaluating risk factors for lameness or a lameness-related condition (other than FMD caused lameness) in sheep raised for the production of meat, milk, or wool in multiple flocks?

- Yes - multiple flocks (proceed to next question)
- No (exclude)

Question 2: Is the full text available in English?

- Yes (proceed to next question)
- No (exclude)
- Not discernible (proceed to next question)

The information to answer Question 2 was contained within the meta-data of the citation. The full text was not retrieved to answer this question.

Question 3: Was the study conducted in a target country?

- Yes (proceed to next level)
- No (exclude)
- Not discernible (proceed to next level)

For citations that proceeded to Level 2 of screening, the following questions were evaluated. The full text was obtained to answer these questions. References for which the full text could not be obtained were excluded and a list of those references has been included in the Results section.

Question 4: Is the full text language English AND the study location a target country?

- Yes (proceed to next question)
- No (exclude)
- Study location not discernible (exclude)

Question 5: Is management system (indoor/outdoor/mixed) a studied risk factor for lameness?

- Yes (proceed to next question)
- No (exclude)

Question 6: Is there more than one study unit per exposure category?

- Yes (proceed to data extraction)
- No (exclude)

Studies that remained after Question 5 were considered eligible for the review and their citations have been provided in an EndNote library.

Appendix F. Manuscript level questions for data extraction and clarification notes for systematic review

These notes provide clarification of decisions about the interpretation of extracted data.

- Study design (retrospective cohort, prospective cohort, case control, cross-sectional/survey, other observational study)
- Year the study was performed. If the study extended over multiple years, the year in which it started was extracted. If a survey was conducted asking the respondent for retrospective data, the year the survey was conducted was extracted, not the earliest year for which the respondent reported. If not reported, the year the study was published was extracted.

- Country (only studies with discernible countries were included in the screening)
- Sheep production type (meat, wool, milk, mixed, not reported). Any studies that discussed milking sheep or milking equipment were classified as milk production studies. Studies that discussed only mastitis were not included as milk production unless they met the above criteria. Studies that reported carcass characteristics were classified as meat production studies. Production of lambs was not considered evidence of a meat production system. Studies that reported wool quality were classified as wool production studies. Shearing sheep was not an indicator of wool production.
- Class of sheep studied for the estimate of lameness. More than one could be selected. (ewes, lambs, rams, whethers, hoggets, flocks, not described)
- Breed of sheep enrolled. The main breed of a cross was selected for crossbreeds (e.g., for “Suffolk-cross,” “Suffolk” was selected) (Scottish Blackface, Welsh Mountain Sheep, Merino breeds, etc.—this list was added to as the reviewers identified more breeds)
- Time of year in which the data was collected. More than one could be selected. (12 months and 4 seasons were possible response options)
- Sampling method of farms/flocks (census of all farms in an area, random from sampling frame, convenience, stratified or random sample from dominant clusters after k-means cluster analysis, not described)
- Total number of farms/flocks enrolled in the study
- Sampling method of animals within farms/flocks (census of all animals on farm, random sample of animals on farm, convenience sample of animals of farm, not described)
- Total number of animals enrolled in study, if available

Question Text	Type	Response options
What is the study design?	Radio	Retrospective cohort
		Prospective cohort
		Case-control
		Cross-sectional/survey
		Other observational design
Year of study? If study extended over multiple years, in what year did it start?	Text	
Country of study?	Text	
Production type?	Checkbox	Meat
		Milk
		Wool
		Mixed
		Not Reported
Class of sheep studied for the lameness estimate. Check all that apply.	Checkbox	Ewes
		Lambs
		Rams
		Whethers
		Hoggets

		Flocks
		Not described
Breed of sheep enrolled. Check off the main breed for crossbreeds (e.g., "Suffolk-cross" would be "Suffolk"). Check all that apply.	Checkbox	Scottish Blackface
		Welsh Mountain Sheep
		Merino breeds
		Texel
		Mule
		Suffolk
		Swaledale
		Polled Dorset
		Chios-breed dairy
		Rygia sheep
		Norwegian pelt sheep
		Bergamasca
		Sarda
		Massese
		Padovana
		Lleyn
		Beltex
		Not described
Time of year in which data were collected. Check all that apply.	Checkbox	Full year
		"Spring"
		"Summer"
		"Fall/Autumn"
		"Winter"
		January
		February
		March
		April
		May
		June
		July
		August
		September
		October
		November
		December
		Not discernible
Sampling method of farms/flocks	Radio	Census of all farms in an area
		Random from sampling frame
		Convenience
		N/A - animal-level study with one flock
		Not described
Total number of farms/flocks enrolled in the study	Text	
Sampling method of animals within farms/flocks	Radio	Census of all animals on farm
		Random sample of animals on farm
		Convenience sample of animals on farm
		Not described

Total number of animals enrolled in the study, if available.	Text	
Notes	Text	

Appendix G. Outcome and exposure level questions and notes for data extraction for systematic review

Outcome information

- Condition assessed. If multiple conditions were included in a single outcome definition, all response options that applied were selected. (lameness/locomotion, footrot, interdigital dermatitis, contagious ovine digital dermatitis, other lameness-related condition)
- The unit of analysis for the outcome. (flock, animal)
- The full definition of the outcome. This was copied directly from the full paper. (e.g., “>5% prevalence of lameness in flock: Y/N”)
- Outcome type (prevalent cases, incident cases)
- Outcome measurement methods (researcher observed, producer reported, not discernible or not reported)

Exposure information

- The full definition of the management exposure. This was copied directly from the full paper. (e.g., “Number of months of the year housed: 0, 1–2, >=3;” “Access to pasture: Y/N”)
- For continuous outcomes:
 - The central tendency metric reported (mean, median, group comparison mean difference, standard mean difference, not reported)
 - For each level of the exposure variable, the central tendency value (copied from text)
 - The dispersion metric reported (SD, SEM, 95% CI, range, interquartile range, other, not reported)
 - For each level of the exposure variable, the dispersion value (copied from text)
- For categorical outcomes:
 - For each level of the exposure variable, the number of animals with the outcome (r). If “r” was different for univariable and multivariable analyses, both numbers were reported.
 - For each level of the exposure variable, the total number of animals assessed for the outcome in the exposure group (n). If “n” was different for univariable and multivariable analyses, both numbers were reported.

Comparison of exposure groups

- The approach used to determine the association between exposure and outcome. All that applied were selected. (univariable regression or other univariable test (e.g., Chi-squared test), multivariable regression, not reported, not discernible)
- For each method of comparison reported (i.e., univariable, multivariable, undetermined), the following data were extracted:
 - The number of units of concern in the analysis (number or not reported)
 - The measure of association. For categorical exposure variables, measures of association were reported for each level, with the referent level = 1. (number or not reported)
 - The 95% confidence interval of the measure of association. For categorical exposure variables, 95% confidence intervals were reported for each level, with the referent level = NA. (number or not reported)
 - The p-value of the measure of association. For categorical exposure variables, p-values were reported for each level, with the referent level = NA, as well as a global p-value. (number or not reported)
 - The point estimate (coefficient). For categorical exposure variables, point estimates were reported for each level, with the referent level = NA.
 - The SE of the estimate. For categorical exposure variables, SEs of the estimates were reported for each level, with the referent level = NA.
- For multivariable methods, the following data were also extracted:
 - The covariates included in the final model (list or not reported)

Question Text	Type	Response options
For this form, what outcome was assessed? If multiple conditions were included in a single outcome definition, choose all that apply. (Repeat this form for multiple outcomes.)	Checkbox	Lameness/locomotion
		Footrot
		Interdigital dermatitis
		Contagious ovine digital dermatitis
		Other lameness-related condition
What was the unit of concern for this outcome?	Radio	Flock
		Animal
What was the full definition of the outcome? (e.g., ">5% prevalence of lameness in flock: y/n")	Text	
What was the definition of the management exposure variable? (e.g., "Number of months of the year	Text	

housed: 0, 1-2, >=3"; "Access to pasture: y/n")		
Was the outcome continuous or categorical?	Radio	Continuous
		Categorical
What was the central tendency metric?	Radio	Mean
		Median
		Group comparison mean difference
		Standard mean difference
		Not reported
For each level of the independent variable, what was the central tendency value?	Text	
What was the dispersion metric?	Radio	SD
		SEM
		95% CI
		Range
		Interquartile range
		Other
		Not reported
For each level of the independent variable, what was the dispersion value?	Text	
For each level of the independent variable, what was "r" (proportion of animals with the outcome)? If "r" was different for univariable and multivariable analyses, report both sets of numbers. (# or not reported)	Text	
For each level of the independent variable, what was the "n" (total number assessed for outcome in the group)? If "n" was different for univariable and multivariable analyses, report both sets of numbers. (# or not reported)	Text	
What was the approach to determining the association between exposure and outcome? Choose all that apply.	Checkbox	Univariable regression modeling or other univariable test (e.g., Chi-square test)
		Multivariable regression modeling
		Not reported
		Not discernible
Univariable: Number of units of concern in the analysis (# or not reported)	Text	
Univariable: What was the measure of association, if reported? For a categorical independent variable, report all measures of association, with referent = 1.	Text	
Univariable: What was the 95% confidence interval of the measure(s) of association, if reported? For a categorical independent variable, report the CI for each level, with	Text	

referent = NA.		
Univariable: What was the p-value of the measure of association, if reported? For a categorical independent variable, report p-value for each level, with referent = NA, as well as a global p-value.	Text	
Univariable: What was the estimate (coefficient), if reported? For a categorical independent variable, report the estimate for each level.	Text	
Univariable: What was the SE of the estimate, if reported? For a categorical independent variable, report the SE of the estimate for each level.	Text	
Multivariable: Number of units of concern in the analysis (# or not reported)	Text	
Multivariable: What was the measure of association, if reported? For a categorical independent variable, report all measures of association, with referent = 1.	Text	
Multivariable: What was the 95% confidence interval of the measure(s) of association, if reported? For a categorical independent variable, report the CI for each level, with referent = NA.	Text	
Multivariable: What was the p-value of the measure of association, if reported? For a categorical independent variable, report p-value for each level, with referent = NA, as well as a global p-value.	Text	
Multivariable: What was the estimate (coefficient), if reported? For a categorical independent variable, report the estimate for each level.	Text	
Multivariable: What was the SE of the estimate, if reported? For a categorical independent variable, report the SE of the estimate for each level.	Text	
Uni/multivariable: Number of units of concern in the analysis (# or not reported)	Text	
Uni/multivariable not discernible: What was the measure of association, if reported? For a categorical independent variable, report all measures of association, with referent = 1.	Text	
Uni/multivariable not discernible: What was the 95% confidence interval of the measure(s) of association, if	Text	

reported? For a categorical independent variable, report the CI for each level, with referent = NA.		
Uni/multivariable: What was the p-value of the measure of association, if reported? For a categorical independent variable, report p-value for each level, with referent = NA, as well as a global p-value.	Text	
Uni/multivariable: What was the estimate (coefficient), if reported? For a categorical independent variable, report the estimate for each level.	Text	
Uni/multivariable: What was the SE of the estimate, if reported? For a categorical independent variable, report the SE of the estimate for each level.	Text	
What covariates were included in the final model? (list or not reported)	Text	
Was the outcome assessment blinded?	Radio	Yes
		No
		Not reported
What was the risk of bias in the outcome assessment in unblinded studies?	Radio	High
		Low
		Unclear
Who measured the outcome?	Radio	Farmer/flock owner
		Researcher
		Not discernible
Notes	Text	

Appendix H. Risk of bias form and interpretation notes for systematic review

The notes for interpreting these questions are available publically available at <http://www.ncbi.nlm.nih.gov/books/NBK154461/>.

Question Text	Type	Answer Text
Do the inclusion/exclusion criteria vary across the comparison groups of the study?	Radio	Yes, they vary
		Partially: some, but not all, criteria applied to all groups or not clearly stated if some criteria are applied to all groups
		No, they do not vary
		Cannot determine: article does not specify
		Not applicable: study has only one group and so does not include comparison groups
Notes about Q1	Text	
Does the strategy for recruiting participants into the study differ across groups?	Radio	Yes, differs

		No, does not differ
		Not discernible
		Not applicable: one study group
Notes about Q2	Text	
Is the selection of the comparison group inappropriate, after taking into account feasibility and ethical considerations?	Radio	Yes, inappropriate
		No, not inappropriate
		Cannot determine or no description of the derivation of the comparison group
		Not applicable: study does not include a comparison group (case series, one study group)
Notes about Q3	Text	
Does the study fail to account for important variations in the execution of the study from the proposed protocol?	Radio	Yes, fails to account
		Partially fails to account
		No, does not fail to account
		Cannot determine
		Not applicable: not an intervention study or no variations
Notes about Q4	Text	
Was the outcome assessor not blinded to the intervention or exposure status of participants?	Radio	Yes, not blinded
		No, blinded
		Not applicable, assessor cannot be blinded
Notes about Q5	Text	
Were valid and reliable measures, implemented consistently across all study participants used to assess inclusion/exclusion criteria, intervention/exposure outcomes, participant health benefits and harms, and confounding?	Radio	Yes, valid and reliable measure used
		No, valid and reliable measure not used
		Cannot determine or measurement approach not reported
Notes about Q6	Text	
Was the length of follow-up different across study groups?	Radio	Yes, different or cannot determine
		No, not different or remedied through analysis
		Not applicable: cross-sectional or only one group followed over time
Notes about Q7	Text	
In cases of high loss to follow-up (or differential loss to follow-up), was the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Radio	Yes, impact assessed

		No, impact not assessed
		Cannot determine
		Not applicable: no loss to follow-up or loss to follow-up was not considered to be high, cross-sectional study, or case-control study selected on outcome
Notes about Q8	Text	
Are any important primary outcomes missing from the results?	Radio	Yes, important outcome(s) missing
		No, important outcomes not missing
		Cannot determine
Notes about Q9	Text	
Are any important harms or adverse events that may be a consequence of the intervention/exposure missing from the results?	Radio	Yes, important harms/adverse events (outcomes) missing
		No, important harms/adverse events (outcomes) not missing
		Assessment of harms not applicable to this study
Notes about Q10	Text	
Are results believable taking study limitations into consideration? This question is intended to capture the overall quality of the study. Consider issues that may limit your ability to interpret the results of the study. Review responses to earlier questions for specific criteria.	Radio	Yes, believable
		No, not believable
Notes about Q11	Text	
Was there any attempt to balance the allocation between the groups or match groups (e.g., through stratification, matching, propensity scores)?	Radio	Yes or study accounts for imbalance between groups through a post hoc approach such as multivariate analysis
		No or cannot determine
		Not applicable: study does not include a comparison group (case series or one study group)
Notes about Q12	Text	
Were important confounding variables not taken into account in the design and/or analysis (e.g., through matching, stratification, interaction terms, multivariate analysis, or other statistical adjustment such as instrumental variables)?	Radio	Yes, not accounted for or identified
		Partially: some variables taken into account or adjustment achieved to some extent
		No: taken into account
		Cannot determine
Notes about Q13	Text	

Comments:	Text	
-----------	------	--

Appendix I. Proposed protocol for review of the literature evaluating the effect of management system on lameness in sheep raised for the production of meat, milk, or wool in Europe

PECO(S) Question (PRIMSA ITEM 4)

In sheep raised for the production of meat, milk, or wool on commercial farms located in Europe (defined as member countries of the OIE European Region) (P), what is the impact of extensive/outdoor/migratory management (E) compared to intensive/indoor management (C) on the farm-level prevalence or incidence of lameness, footrot, and other lameness-related conditions measured during non-outbreak periods (O) in observational studies or experimental studies with management type as the exposure of interest allocated at the farm-level (S)?

Eligibility criteria (PRIMSA ITEM 6)

Relevant participants

The relevant study population is defined as sheep raised for the commercial production of meat, milk, or wool in the European region. The European region is defined as member countries of the OIE European region (http://www.oie.int/fileadmin/Home/eng/About_us/docs/pdf/2009_Commission_Europe_A.pdf), and includes northern Asia (e.g., Russia, Ukraine, Uzbekistan), Turkey, Israel, and several other non-EU countries.

Relevant exposure or comparisons

Relevant studies must report the management system type under which the sheep were raised (e.g., outdoors, indoors, or mixed (both indoors and outdoors)). Management system is defined as the type of housing as it relates to exposure to the environment (i.e., indoors or outdoors). Management system is not defined as animal density, a management factor. Rotational grazing is an outdoor management system, with high stocking density. Traditional grazing is an outdoor management system with lower stock density. Outdoor housing is also not defined as an absence of structures. For example, feedlot animals may be raised in an outdoor management system but confined in pens and animals in extensive outdoor grazing systems may be offered shade structures. These latter two examples would still be defined as outdoor systems. Systems that house animals for management practices, such as housing for lambing or shearing, but otherwise place animals on pasture would be defined as outdoor systems. Finally, outdoor management is not defined as access to pasture, as feedlot animals may not have access to pasture but are raised outdoors.

Types of outcome measures

Outcomes of interest are measures of lameness and lameness-related conditions. Lameness-related conditions include (but are not limited to) footrot, contagious ovine digital dermatitis, interdigital dermatitis, scald, white-line disease. As the exposure occurs at the farm level, the outcome measure must also be the farm level. Measurement may be made by either the researcher or the producer. There are numerous approaches possible for researchers to summarize each farm-level outcome. Researchers may measure disease frequency at the individual level to calculate the farm-level incidence or prevalence or the farm-level presence or absence of lameness at a given prevalence or incidence threshold. We will not restrict studies based on the approach to describing lameness at the farm level.

Incidence may be measured as incidence rate or incidence risk. The approach to comparison of management systems would be the comparison of the average incidence or prevalence of lameness on exposed farms versus unexposed farms. Possible approaches to compare measures of lameness across management systems will vary by researcher. We will not use the approach to summarizing the association as an exclusion criterion.

Relevant study designs

Observational studies relevant to the review are cohort, case-control, or cross-sectional studies that measure the association between management types (e.g., extensive/outdoor/migratory and intensive/indoor systems) and farm-level incidence or prevalence of lameness, footrot, or another lameness-related condition. Studies will be required to have more than one observation for exposures of interest (i.e., more than one farm per exposure level). Observational studies that compare the prevalence of lameness on one farm with a single type of management system with the prevalence of lameness on another single farm with a different type of management system will not be included due to confounding by other practices between farms. Although the association between outcome and the exposure of interest is likely confounded by other factors, we will not exclude studies that only conducted univariable analyses.

Information Sources (PRIMS ITEM 7)

Searches of the electronic citations databases listed in the table below will be conducted. In addition to searches of published literature in bibliographic databases, we will also undertake searches of two conferences. A key conference of interest is the International Workshop on Assessment of Animal Welfare at Farm and Group Level. This conference has been hosted 5 times. ISU has access to the 1st conference proceedings (on-site copy of conference proceedings), the 2nd–4th conference proceedings were published in the journal *Animal Welfare*, which is indexed in CABI. It is currently unclear if the 5th conference held in Guelph has been published in *Animal Welfare*; however, the proceedings are available online at <http://www.uoguelph.ca/csaw/wafl/documents/WAFLproceedingsweb.pdf>. We will also attempt to find the conference proceedings for the International Conference of Lameness in Ruminants. This conference has been hosted nine times, most recently in 2013 (<http://www.bristol.ac.uk/vetscience/lamenessconf/>). To date, we have not been able to identify a location for these conference proceedings. They are not currently held at ISU library. Unless they are available electronically or quickly through interlibrary loan, it may not be possible to search these proceedings in the first month of the contract. We will need to complete data extraction with 3–4 weeks of the contract start date, if we are to finish the analysis. As these conference proceedings contain only abstracts with little information, we will identify potentially relevant conference abstracts, then search for the title, first author, and last author in CABI to find full publications. We will not extract data from conference papers with fewer than 500 words. The reference lists of relevant studies will also be hand-searched for additional evidence.

Electronic citation databases to be searched for the review

Information resources	Interface
Science Citation Index (SCI)	Web of Knowledge, Thompson Reuters
Conference Proceedings Citation Index – Science (CPCI-S)	Web of Knowledge, Thompson Reuters
CAB Abstracts	Web of Knowledge, Thompson Reuters
BIOSIS Previews	Web of Knowledge, Thompson Reuters
MEDLINE and MEDLINE In-Process	OvidSP
AGRIS	http://agris.fao.org/
AGRICOLA	http://agricola.nal.usda.gov/

TEKTRAN	www.ars.usda.gov/services/tektran.htm
CRIS	http://cris.nifa.usda.gov/
Science.gov	www.science.gov/
ScienceResearch.com	http://scienceresearch.com/scienceresearch/
Open Grey	www.opengrey.eu/

Search strategy (PRIMSA ITEM 8)

The search strategy comprises two key elements: a) sheep (search line 1) and b) lameness or conditions commonly resulting in lameness in sheep (search lines 2 and 3). In order to ensure the search is adequately sensitive, the key concept of prevalence will not be built into the search strategy. This concept is often poorly or inconsistently described in the title, abstract, and indexing of database records and, therefore, its inclusion may result in relevant studies being missed. An example search is listed in Figure 1.

The searches will not be limited by date or language. Where the functionality of the database interface allows, the searches will be limited to try to restrict the results to those related to a European context. This will be achieved using the “NOT” operator to remove records that include the name of a North American, South American, African or Asian country in the title, abstract, or author affiliation field, **but not** the name of countries belonging to the European Regional Commission of the World Organisation for Animal Health (OIE).

The search strategy will be adapted appropriately to run in each information source. Adaptation includes consideration of database interface differences as well as adaptation to different indexing languages.

Figure 1: Draft search strategy to identify evidence on sheep lameness in Science Citation Index

1. TS=("sheep" OR "ovine" OR "lamb" OR "lambs" OR "ewe" OR "ewes" OR "ram" OR "rams" OR "mutton" OR "hogget*" OR wether*)
2. TS=("lame" OR "lameness")
3. TS=("footrot*" OR "foot rot" OR "overgrown hoof" OR "overgrown hooves" OR "over-grown hoof" OR "over-grown hooves" OR "pododermatitis" OR "podo-dermatitis" OR "scald" OR "digital dermatitis" OR "interdigital dermatitis" OR "CODD" OR "soil balling" OR "fibroma" OR "granuloma" OR ("foot" OR "feet" OR "toe" OR "toes" OR "hoof" OR "hooves" OR "white line" OR "whiteline") NEAR/2 ("abscess*" OR "infect*"))
4. 1 AND (2 OR 3)

Key:

TS: topic field search

near/n: proximity search

*: truncation

Search considerations

The timeframe for the review is two months. Therefore, translations will not be conducted for papers that are not available in English. We will include in the review papers that could be obtained within 1 month after the start of the contract to enable the team to complete the remaining aspects of the

review. Relevant papers identified but not obtained within that time frame will be indicated in the final review report.

Search results and analysis

The results of the searches will be downloaded into bibliographic management software (EndNote 7) and de-duplicated using several algorithms before uploading to DistillerSR® (Evidence Partners©, Canada, 2012), an internet-based systematic review software, for relevance screening, data extraction, and management of identified studies.

Study Selection (PRIMSA ITEM 9)

Two reviewers will independently perform the relevance screening exercise, with disagreements on the inclusion or exclusion of studies resolved by consensus or with the input of the review coordinator.

Screening for eligibility for the review

There will be two levels of screening. The rationale for two levels is that some questions will likely require the full paper to evaluate. The following questions will be used to determine whether a study will be included in the review.

Question 1: Does the citation describe an observational study evaluating risk factors for lameness or a lameness-related condition in sheep raised for the production of meat, milk, or wool?

- Yes (proceed to next question)
- No (exclude)

If “Yes” is selected in Question 1, the following question will be asked:

Question 2: Is the full text available in English?

- Yes (proceed to next level)
- No (exclude)
- Not discernible (proceed to next level)

The information to answer Question 2 will be contained within the meta-data of the citation. The full text will not be retrieved to answer this question. For citations that proceed to Level 2 of screening, the following questions will be evaluated. If the required information cannot be obtained from the abstract, the full text will be obtained to answer these questions.

Question 3: Was the study conducted in a target country?

- Yes (proceed)
- No (exclude)

Question 4: Is management system (indoor/outdoor/mixed) a studied risk factor for lameness?

- Yes (proceed)
- No (exclude)

Question 5: Is there more than one study unit per exposure category?

- Yes (proceed)
- No (exclude)

Studies remaining after Question 5 will be considered eligible for the review. Citations that identify these studies will be provided in an EndNote library. A flow diagram will indicate studies that were eligible for the review but could not be obtained during the contract period.

Screening for eligibility for meta-analysis

If a meta-analysis is possible, studies may be excluded from the meta-analysis if reporting is incomplete. A flow diagram will indicate which studies were eligible for the review but excluded from the meta-analysis due to incomplete reporting.

Data collection process (PRIMSA ITEM 10)

Two reviewers will extract data independently from studies deemed to be relevant to the review. Data will be compared between the reviewers and any conflicts will be resolved through discussion. Data extraction forms will be designed in DistillerSR®. Initial forms will be designed and piloted on several papers and modified as required for use.

Data Items (PRIMSA ITEM 11)

Information collected from each study will consist of, but is not limited to the following:

General study characteristics:

- Study design (retrospective cohort, prospective cohort, case control, cross-sectional/survey)
- Year the study was performed (if not reported, will use the year the study was published)
- Country (must be reported or study will be excluded)
- Sheep production type (meat, wool, milk, mixed, not reported). We will classify any studies that discuss milking sheep or milking equipment as milk production studies. Studies that discuss only mastitis will not be included as milk production unless they meet the above criteria. Studies that report carcass characteristics will be classified as meat production studies. Production of lambs is not considered evidence of a meat production system. Studies that report wool quality will be classified as wool production studies. Shearing sheep will not be an indicator of wool production.
- Total number of farms/flocks enrolled in the study
- Class of sheep studied for the estimate of lameness (flocks, ewes only, lambs only, rams only, not described)

- Breed of sheep enrolled (Scottish Blackface sheep, Welsh Mountain Sheep, etc.—this list will be added to as the review identifies breeds)
- Time of year in which the data was collected (month(s) or season(s), as reported—Check all that apply)
- Sampling method of farms/flocks (random from sampling frame, convenience, not described)
- Sampling method of animals within farms (census of all animals on farm, random sample of animals on farm, convenience sample of animals of farm, not described)

Outcome information:

- Description of outcome assessed (lameness, footrot, contagious ovine digital dermatitis, white line disease, scald, over-grown claws, laminitis, all causes of lameness, other)
- Outcome type (prevalent cases, incident cases)
- Outcome measurement methods (researcher observed, producer reported, not discernible or not reported)

Study cohort-level information

The outcome will be identified first, then the exposure data for each outcome will be collected.

- **Exposure information** (1st group, 2nd group, 3rd group)
 - Type of management system (extensive pasture, intensive pasture, outdoor penned, housed entirely indoors (artificial/wood/cement floors), housed indoors with natural floors, mixed housing (indoors some times of the year))
 - Unit of concern (whole farm, subgroup on farm, not discernible)
 - Outcome metric reported
 - Number of farms with the outcome in the exposure category (r)
 - Average prevalence in exposure category
 - Median prevalence in exposure category
 - Range of prevalence in exposure category
 - Number of farms in exposure category (n)

Approach to comparison of groups

- Metric used to compare groups

- Mean difference (average prevalence in exposed group - average prevalence in unexposed group)
- Risk ratio
- Odds ratio
- Prevalence ratio
- Hazard ratio
- Approach to estimating the metric used to compare groups
 - Adjusted estimate (adjusted for covariates but no explicit adjustment for confounding). This category includes any studies using step-wise algorithms for model building. Check all that apply. We will add to this list of covariates as analysis proceeds:
 - A measure of foot trimming
 - A measure of season
 - A measure of genetics (includes breed)
 - Confounder-adjusted estimate (explicit adjustment for confounding) Check all that apply. Will add to the above list as analysis proceeds.
 - Unadjusted estimate (univariable regression, t-test, or chi-squared test, or non-parametric equivalents if used a category definition of lameness (i.e., a presence/absence over a threshold))
- Value of metric (numerical, blank if not reported)
- Metric used to reported variation
 - Confidence interval
 - Standard error of metric
- Value of the reported variation (numerical, blank if not reported)
- Metric used to assess probability of the null hypothesis
 - Exact P-value
 - Significance testing (significant/not significant or $p < 0.05$)
- Value of metric used to assess probability of the null hypothesis (numerical, blank if not reported)

Assessment of risk bias in included studies (PRISMA ITEM 12)

No standard tool is available for the assessment of the risk of bias for observational studies. Instead we propose to assess studies for the following potential sources of bias:

- Information bias for the exposure — Was the extent of information bias on the exposure variable likely to be non-differential? (e.g., were different evaluations of the exposure applied to different groups?) (yes = high risk of bias, no = low risk of bias, unclear risk of bias)
- Selection bias — Was the approach to enrolment likely to be associated with differential selection probabilities for different outcome groups? (e.g., farm with indoor management systems with high prevalence of lameness were more likely to be enrolled than other groups) (yes = high risk of bias, no = low risk of bias, unclear risk of bias)
- Confounding — Were known confounders identified *a priori* and controlled for, either by restriction, matching, or multivariable analysis? (yes = lower risk of bias, no = high risk of bias)

Summary measures (PRISMA ITEM 13)

Often systematic reviews use the individual as the unit of concern; therefore, summary measures of effect can be risk ratios or odds ratios for dichotomous outcomes, or mean differences for continuous outcomes. When a group-level metric such as farm-level prevalence or incidence is used as the outcome in the primary study, the possible approaches to analysis are extremely varied. Some studies may calculate the average prevalence in exposed and unexposed groups and report the mean difference in that average. Other studies may conduct a multilevel regression analysis, with individual lameness as the outcome and farm-level variables that do not vary between animals on the farm as covariates, including farm as a random effect and management type as a fixed effect. In this situation the outcome could be an odds ratio or hazard ratio. Because of the potential heterogeneity of analyses across studies, at this point it is not clear if it will be possible to calculate a summary measure across all studies. Sometimes it is possible to use a standardized mean difference as a summary metric, when all the studies measure a similar but not the same metric. However, in these situations the data need to be of the same form (i.e., all continuous outcomes). It is unlikely that calculation of a summary measure will be an option for this review.

Dealing with missing data

Given the time frame, we will not contact authors to obtain missing data. This is a potential limitation of the review. Recently we conducted a review with continuous outcomes such as would be considered in this review. Around 30% of original papers did not report measures of variation and by contacting the authors, we were able to obtain information on numerous papers. However, this was a long process (months). Imputation methods for studies that do not report measures of variation for the outcomes of interest will not be used.

Assessment of heterogeneity

We will conduct assessments of heterogeneity visually and using statistical approaches, if appropriate. With respect to clinical and methodological sources of heterogeneity, we will conduct formal statistical assessment of heterogeneity where at least 10 studies have reported the characteristics for any one outcome. We will assess overall heterogeneity using chi-squared tests and I^2 . As appropriate, we will conduct subgroup analysis for covariates identified in the study-level characteristics and meta-regression. We propose, if the sample size is sufficient, to conduct a meta-regression to determine what factors are associated with the magnitude of effect. In this analysis we will determine if putative sources of methodological heterogeneity, such as the approach to controlling confounders, and

putative sources of contextual heterogeneity are associated with different effect sizes. We will assess the validity of the model before reporting any outcomes.

Data synthesis

The approach to evidence synthesis will depend upon the frequency of the outcomes of interest within the relevant studies. Tables that describe the outcomes used and the associations observed will be reported. We will attempt to prepare forest plots that compare the associations. For example, if several studies report an odds ratio for the association between management systems (indoor/outdoor and mixed) and lameness (using some categorical definition), then those will be plotted on a series of pairwise forest plots (outdoors versus indoors, outdoors versus mixed, indoors versus mixed). Similarly, if several studies report mean differences of the prevalence of lameness between management systems, we may plot all possible pairwise comparisons. We will limit the presentation of forest plots to outcomes of interest with at least two independent results. We will only present summary effect sizes for outcomes where we are able to assess heterogeneity, otherwise only the individual studies' results will be plotted. The rationale is that a single evaluation of an outcome provides little evidence of impact, if it has not been replicated at least once. As discussed above, determination of a summary effect size is an unlikely product of the meta-analysis. In our experience, when observational studies are reported, the confounders that are controlled are different between studies, and combining the different study estimates is often inappropriate, similar to combining apples and oranges. A summary effect measure based on observational data has a higher potential to be biased and often a more appropriate approach to describing the data is to report the individual study results in a manner that enables others to observe the effect and discern if an effect seems consistent between studies.

Approach to presenting the results

Study selection (PRISMA ITEM 17)

We will use a flow chart as recommended by PRISMA to present the number of papers screened, the number of relevant papers, and the number of papers included in the meta-analyses (if conducted)

Study characteristics (PRISMA ITEM 18)

We will provide a table that contains information about the relevant studies and other general characteristics collected and discussed in section 6.11.1.

Risk of bias within studies (PRISMA ITEM 19)

We will provide a table that contains this information about relevant studies.

Results of individual studies (PRISMA ITEM 20)

We will provide a table that contains this information about relevant studies. It is possible that there will be several tables, given the potential variety of outcomes. For example, if numerous studies are found that assess lameness and another set that assess foot rot, it may be better to provide two sets of tables, although some duplication will occur. If suitable, we will provide a forest plot(s) that contains individual study data in lieu of a table.

Synthesis of results (PRISMA ITEM 20)

If a synthesis is possible, we will provide a meta-analysis with a forest plot and the summary effect measure included. We will also provide the metrics for heterogeneity assessed. If only a narrative report is possible because the heterogeneity of results suggest synthesis is not suitable, this will be provided with the test statistics and p-values.

Risk of bias across studies (PRISMA ITEM 21)

If an analysis to assess small study effects is possible, we will provide the results of that analysis. If not, we will comment on the potential for small study effects.