Report on Animal Welfare of Dairy Cows in Indoor Loose Range Housing: A Systematic Review

18th November 2022 - Final Report

About OXSREV

Oxford Systematic Reviews (OXSREV®) are specialists in evidence-based solutions. We specialise in stakeholder engagement/workshops, online surveys, high impact evidence evaluation reports, systematic reviews, systematic maps, rapid reviews, peer reviewed journal articles, and interactive location-based knowledge maps.

The Project Team

The project team at OXSREV® comprises Dr. William J. Harvey[1,2] (Managing/Founding Partner), Dr. Gillian Petrokofsky[1,3] (Founding Partner), Mr. Leo Petrokofsky[1] (Founding Partner), Mr. Matthew Jordon[4] (Consultant), Mr. Ali Malik[1] (Consultant), Mr. Thomas Carter[1] (Consultant), and Mr. Luke S. Wade[1] (Consultant). This team has successfully worked together over many years, initially in the Long-term Ecology lab of University of Oxford, and more recently in OXSREV® to provide services to governments, NGOs, Universities, as well as to the private sector.

[1]Oxford Systematic Reviews, 266 Banbury Road, Oxford, OX2 7DL, UK
[2]School of Archaeology, University of Oxford, 36 Beaumont Street, Oxford OX1 2PG, UK
[3]Department of Biology, University of Oxford, Mansfield Road, Oxford, OX1 3SZ, UK
[4]Oxford Martin School, University of Oxford, 34 Broad Street, Oxford, OX1 3BD, UK

[*]Corresponding author: william.j.harvey@oxsrev.org

Citation

Harvey, W.J., Petrokofsky, L., Malik, A., Carter, T., Wade, L.S., Jordon, M., & Petrokofsky, G. 2022. Report on Animal Welfare of Dairy Cows in Indoor Loose Range Housing: A Systematic Review. Oxford. 46 pp.

Acronyms

CCA - Canonical Correspondence Analysis
HWB - health, welfare and behaviour
PICO - Population, Intervention, Comparator and Outcomes, pertaining to structure of review question

Executive Summary

Assessing the evidence base systematically as a whole:

- There is only weak evidence that year round loose-housing (zero grazing i.e. with or without outdoor access in yards) is either better or worse than housing systems with grazing for the health and welfare of dairy cows.
- There is also only weak evidence that year-round loose-housing with zero outdoor access is either better or worse than housing systems with any outdoor access, including but not limited to grazing, for the health and welfare of dairy cows.
- There is too much variation across studies in terms of what is measured and how it is measured to allow robust statistical analysis of the direct effects of loose-housing systems and/or grazing on the health and welfare of dairy cows. In addition, data are often presented in an aggregated form that does not allow meaningful comparisons.
- The evidence base contains even less data that allows for comparisons between different design features of loose-housing systems. Therefore meaningful comparisons can not be made.
- There is even less evidence for behaviour indicators and there were none that showed only positive or only negative for year round loose-housing without outdoor access or zero grazing.
- For future research, data collected should be made freely available in a disaggregated form to enable further analysis to be conducted, beyond the studies they are collected for.
- For policy/practice to change on the basis of evidence, there would need to be more primary research studies measuring each health and welfare indicator using standardised methods and metrics.
- The current review provides a framework that can incorporate such new research.

Background

Legislation in Sweden requires dairy cows to be grazed during summer months for a minimum number of days, which vary from south to north of the country. This mandatory grazing period was introduced to protect the health and welfare of dairy cows and to allow them to express natural behaviours. In other countries with similar production systems, year-round loose-housing with no access to grazing is gaining popularity. There are strongly-held views both in favour of retaining the mandate for grazing but also for allowing farmers to choose zero-grazing systems. In addition to health, welfare and behaviour considerations, reasons for and against any change in legislation include social considerations of the agricultural landscape. There is a large volume of global research on the health and welfare implications of raising dairy cows in different systems, but relatively few that directly compare year-round loose-housing systems with systems that allow grazing or outdoor access. In order to contribute to evidence-informed policy, a systematic evidence evaluation was undertaken to determine what the objective scientific evidence base is for this question, and to assess whether the data in the evidence base are robust enough to determine any association between housing systems and health, welfare and natural behaviour in dairy cows.

Method

Systematic evidence evaluation was developed in medicine as a rigorous, transparent and repeatable literature synthesis technique that reduces review bias. Typically systematic reviews are conducted in collaboration with stakeholders from academic research, policy and practice and often use a ‘PICO’ framework (defining Population, Intervention, Comparator and Outcomes for the review question). The current review collaboratively developed a strong and rigorous methodological framework (the Protocol) to describe how articles were searched, selected and synthesised. The literature search was conducted in six bibliographic databases to find academic articles, and also embraced a wide search for reports outside the academic literature (grey literature). Clear inclusion criteria were developed to ensure that articles assessed dairy cows in temperate

regions (the ‘population’) and compared a year-round loose-housing system (the ‘intervention’) with another system of housing (the ‘comparator’). To be included, the articles had to report data for health, welfare and behaviour (HWB) of cows in the compared housing systems; articles which only discussed these issues without primary data were not included. Articles which provided data for multiple interventions were treated as separate ‘studies’ and all comparisons between housing systems were captured within each study separately. All HWB indicators (‘outcome’ measurements) were captured in an extensive data coding and data extraction sheet using a template agreed in the Protocol. Data relating to the housing system were extracted together with details of the methods used, location of, and other factors that could affect comparability of the research.

Results

The evidence base of data curated for this review found 53 articles from 24 countries. There were seven different types of housing systems compared across the evidence base. A total of 120 different HWB outcomes were extracted comprising 839 measurements for HWB. The evidence base is summarised in an interactive geo-based evidence map, and also in a series of tables, figures, and narrative accounts of the included studies. The data-set was analysed using two approaches to reflect the review question: (i) year round loose-housing systems with zero grazing (with or without outdoor yards) vs housing systems with access to grazing; and (ii) year round loose-housing systems without outdoor access vs housing systems with outdoor access (yards or grazing).

The evidence base presents a complex picture. On the basis of reported HWB outcomes, some studies reported benefits of year round loose-housing systems; other studies reported disadvantages of year round loose-housing systems; yet further studies reported no difference between year round loose-housing and other systems for individual health welfare, and/or behaviour outcomes. These differences were not just between studies, there were also differences within-studies for individual HWB indicators.

There was substantial heterogeneity in methods of collecting and measuring HWB outcomes across the studies, therefore a robust statistical test (such as meta-analysis) of correlation between potential explanatory variables and HWB outcomes was not possible for any housing comparison or any individual HWB measurement. However, it was possible to analyse the data set using Canonical Correspondence Analysis (CCA), a widely-used statistical technique that identifies and measures the associations among two sets of variables. Some HWB outcomes were reported by single studies only, while data for other HWB outcomes was available from multiple studies. The data for the CCA were therefore filtered to include only HWB metrics drawn from more than one study.

The results from the CCA examining (i) year round loose-housing systems with zero-grazed vs housing systems with access to grazing; and (ii) year round loose-housing systems without outdoor access vs housing systems with outdoor access explained between 3.5-6.1% of the variation in reported outcomes across the study database. This means that 93.9-96.5% of variation within the data sets are not explained by either a specific housing system or presence/absence of grazing.

The HWB data were tabulated to show the number of incidences (scores) where (i) year round loose-housing with zero-grazing was reported to have a beneficial, negative or no-difference effect compared with systems with grazing; and (ii) year round loose-housing without outdoor access was reported to have a beneficial, negative or no-difference effect compared with a housing system with any form of outdoor access. Again excluding single-study data, one out of the 26 HWB outcomes (**sole lesion**) reported only positive and no-difference results (and the positives outweigh the no-difference, meaning that there were more counts of positive than counts of no-difference)) for zero grazed year round loose-housing. Three HWB outcomes (**interdigital hyperplasia**, **sole haemorrhages**, **and clinical lameness**) for zero grazed year round loose-housing had only negative and no-difference results (and the negatives outweigh the no-difference, as above) compared to systems with grazing. Similarly, **sole lesion** HWB outcome had only positive and no-difference results for housing systems without outdoor access compared with housing with outdoor access.

Three of the 18 HWB filtered outcomes (**sole haemorrhages**, **interdigital hyperplasia**, **locomotion**) reported only negative and no-difference results (and the negatives outweigh the no-difference) for housing systems with outdoor access compared with housing with outdoor access. In summary, only one HWB measurement (**sole lesion**) showed positive trends in year-round loose housing with zero grazing and year-round loose housing systems without outdoor access compared with housing systems with outdoor access and/or grazing. Whereas a total of two indicators (**interdigital hyperplasia** and **sole haemorrhages**) showed negative trends in year-round loose housing with zero grazing and year-round loose housing systems without outdoor access compared with housing systems with outdoor access or grazing. An additional negative trend for **clinical lameness** was found for year-round loose housing systems with zero grazing only, and **locomotion** for year-round loose housing systems without outdoor access. However, it must be noted that sole haemorrhages can be seen as a form of sole lesion but opposing trends have been found for these two outcomes, indicating the extent to which the data available is limited.

These results represent avenues for future exploration, rather than definitive results of benefits or disadvantages of year round loose-housing systems and/or access to grazing because of the limitations of the data discussed above.

Animal Welfare of Dairy Cows in Indoor Loose Range Housing: A Systematic Review

Background

The inspiration for the current review is the question, raised in the farming sector in Sweden, of whether the current mandatory grazing for dairy cows during summer months is in reality associated with a higher animal welfare compared with non-grazing housing, given that dairy farms in different countries are increasingly keeping their cows for longer periods indoors in loose range housing facilities. Or to turn the question around: can year-round loose housing of dairy cows offer the same level of all aspects of animal welfare and health compared with summer grazing cows? The question is at the intersection of the pressure for intensification of livestock production to meet the food demands of people on the planet (Steinfield et al. 2006) and animal welfare concerns, which have shaped modern agricultural legislation since the publication of Ruth Harrison's influential book 'Animal Machines (1964), which highlighted the effects of the increasing industrialisation of agriculture.

The book stimulated a commission of enquiry in the UK on the lives of animals raised in intensive, confined conditions. The resulting Brambell report defined animal welfare as "a wide term that embraces both the physical and mental well-being of the animal. Any attempt to evaluate welfare, therefore, must take into account the scientific evidence available concerning the feelings of animals that can be derived from their structure and functions and also from their behaviour" (HMSO, 1965). The Brambell report had no regulatory authority in the UK, but laid the groundwork for European Union legislation, including the Animal Health Law, which though dealing with transmissible animal diseases, amended and repealed a number of animal health laws (Reg. EU 2016/429) and many individual European countries' legislation, notably the 1988 Swedish law, which severely restricted the confinement of agricultural animals. Astrid Lindgren was influential in developing this law, with fellow Swedish farm animal veterinarian Christina Forslund and the law was named 'Lex Lindgren' in honour of her advocacy.

The Five Freedoms

The concept of Five Freedoms was implicit in the Brambell report and formed the basis of later guidance and legislation worldwide relating to the welfare of farm animals in confinement (FAWC, 2009):

1. Freedom from hunger and thirst- by ready access to water and a diet to maintain full health and vigour.
2. Freedom from discomfort- by providing an appropriate environment, including shelter and a comfortable resting area.
3. Freedom from pain, injury and disease- by prevention or rapid diagnosis and treatment.
4. Freedom to express normal behaviour- by providing sufficient space, proper facilities and company of the animal's own kind.
5. Freedom from fear and distress- by ensuring conditions and treatment, which avoid mental suffering

The five freedoms framework can be used to analyse welfare within any livestock system, and continues to be critiqued (see, for example, Fraser 2008), and developed, for example, the concept of Positive Animal Welfare (PAW) has emerged, which goes beyond the "absence of sufferings" and "absence of negatives" (Lawrence et al. 2019).

There has been steadily increasing pressure on animal agriculture in many countries to regulate production practices to assure or improve animal welfare. Many factors have driven this increased interest, not least the increasing awareness by the general public of the intensification of the dairy industry, although research shows that public opposition to intensive ('factory farm') systems are not directly related to animal welfare. In addition, Robbins et al. (2016) reported in their literature review of 150 papers that there was no relationship between animal welfare and size of farm (mostly dairy farms), despite the fact that many smaller farms rear cows outdoors in extensive systems and tend to use more workers per animal.
Regulation can be voluntary, involuntary or a combination of both (Knierim & Pajor 2018, Vapnek & Chapman 2010). Involuntary regulation- compliance with laws, regulations, and other legal instruments- is the main approach in the European Union (EU), with bans on some production systems and animal welfare standards for rearing, transport and slaughter. In non-EU countries, including the USA, Canada, Australia, and New Zealand, voluntary regulation is the main regulatory system (Whiting 2013, Mench 2008). These can take the form of standards developed by commodity groups and audited by independent bodies to ensure continued compliance, or standards set by food retailers. These retailer-driven standards are also important in individual EU countries, where farmers risk access to large markets if they do not or cannot comply.
It is likely that the trend towards greater pressure to address animal welfare concerns in farming will continue along the lines developed from the Five Freedoms, and set out as General Principles by the World Organisation for Animal Health (Fraser et al. 2013). These concerns will be influenced by three, interrelated, factors — "the rise in demand for animal products in developing countries and the associated globalisation of production and trade in animal products; the importance of minimising the environmental impact of animal agriculture; and reconciling demands for animal products with the public's evolving expectations about the treatment of animals" (Mench 2020).

The public may have little knowledge about animal production practices, however, and their opinions about what constitutes acceptable dairy farming, and in particular access to grazing and pasture, may not align with those of farmers (Smid et al. 2022, van den Pol-van Dasselaar et al. 2020). This lack of agricultural knowledge presents a problem in terms of reconciling producer vs consumer perceptions for evidence-informed policy in the future. Zero-grazing systems, where dairy cows are kept indoors permanently, are common in North America and, increasingly, in some parts of Europe (van Vuuren & van den Pol-van Dasselaar 2006), but are often not popular with the public at large. Van den Pol-van Dasselaar et al. (2020) report that stakeholders considered grazing important, but the general public, as a group, considered it especially important (64% gave it the highest importance, compared with 31% in the farmer group and 46% in the scientist group).

Status of outdoor exposure in Europe

There has been a trend towards less grazing and outdoor access in recent years, and in Europe this trend differs markedly between regions, depending on a number of social, economic and geographical factors (van den Pol-van Dasselaar et al. 2020). They report estimates from the European Grassland Federation between 2010-2019 for different regions of Europe as follows:

Northern Europe (Finland, Norway, Sweden)
Welfare legislation is a driver for cows being outside for six weeks to four months per year. In Sweden all dairy cows are obliged by law to go outside; in Norway and Finland, this is required only for cows in some (mainly tie-stall) systems, representing an estimated 80% in 2019 of all dairy cows. The length of the mandatory summer grazing period depends on the region: the further north, the shorter

the grazing season. However, overall, the number of hours that cattle spend outside has been decreasing in recent years.

Western Europe (Ireland, UK)
The percentage of systems with at least some grazing is high (95-100% in Ireland, 70-80% in the UK). In Ireland, grass based seasonal systems of milk production still predominate. The length of the grass-growing season varies from about eight months in the northeast to up to 11 months in the far southwest.

Central Europe - countries with more than 50% grazing- (Belgium, France, Luxembourg, Netherlands, Switzerland)
The percentage of grazing cows has decreased over time; fewer cows are grazing, and both number of days per year and number of hours per day of grazing have decreased. The introduction of a premium price for milk from cows that graze at least 120 days/year for at least 6 hours/day (e.g. Grazing Charta milk) in the Netherlands, Belgium and France has slowed the decline. All-year loose housing system without access to pasture is not common in Switzerland; most cows regularly go on pasture from April to October because farmers receive direct payments for this.

Central Europe - countries with less than 50% grazing (Austria, Denmark, Germany)
Rapid decline of grazing in the early 2000s with only a minority of the cows grazing by 2019 (44% Austria, 20-25% Denmark, 15-40% Germany), possibly as a result of increase in average herd size and associated decrease in available grazing per cow.
Van den Pol-van Dasselaar et al. report that data in Eastern Europe (Bosnia Herzegovina, Bulgaria, Czech Republic, Estonia, Lithuania, Hungary, Poland, Slovenia) and South Europe (Greece, Italy, Portugal, Spain) are limited, but generally grazing has not been a common practice during the last decade, except in Bulgaria and Lithuania, and farm sizes are increasing, with possible loss of available grazing land, as in Central Europe. In the South, climate is an important factor, with warm and dry summers leading to little grass growth and increased risk of heat stress which encourages adoption of indoor systems.
Understanding the health and welfare implications of reduced exposure to the outdoors is a necessary precursor to proposing changes to legislation, in countries with mandatory grazing periods, and for responding to consumer demands for transparency in livestock production systems.

Sweden's legislative environment

Although the current review is not confined to dairy cow housing in Sweden, it is worth summarising the legislative situation in that country as a way of 'setting the scene' for people elsewhere for whom the 'perceived wisdom' is that cows are healthier and happier when they can graze for all or part of their existence.
Keeping cows on pasture in the summer has been a long-standing tradition and a natural component of the production system. In 1988, the Swedish Parliament introduced an Animal Welfare Act in order to raise the level of animal welfare in Sweden and introduced the term "natural behaviour" in the legislation. The intention was to move animal production to models where the animals' needs were better met and they are allowed to express important behaviours. Examples were given in the Bill to this Act (supporting legal material) of what was intended by "natural behaviour", one of which was cows being allowed to graze. The legal documents specify that keeping cows on pasture in the summer is important for the welfare of cows that are tied up during the winter as well as those kept in loose housing systems. In 2019 a new Animal Welfare Act entered into force and the phrasing concerning "natural behaviour" was further elaborated to clarify existing legislation.

The current legal situation concerning pasture, therefore, is:
(i) Animal Welfare Act (2018:1192)[1] , Chapter 2, Section 2.
Animals shall be kept and cared for in a good environment for animals and in such a way that:

- their well-being is promoted;
- they are able to perform behaviours for which they are strongly motivated for and that are important for their well-being (natural behaviour); and
- behavioural disorders are prevented.

(ii) Animal Welfare Ordinance (2019:66)[2]
Chapter 2, Section 3

> *Cattle kept for milk production that are older than six months shall be kept in pasture in the summer. The first paragraph does not apply to animals being kept in quarantine.*

Chapter 2, Section 4

> *Cattle, other than cattle kept for milk production, shall be kept in pasture or allowed access to outdoor areas in the summer in some other way.*

The first paragraph does not apply to animals younger than six months, bulls or animals being kept in quarantine. Other animal welfare regulations for cattle, available in Swedish only, also cover pasture and grazing requirements.[3]

These regulations together regulate how many days and what period (summer months but part of the spring and autumn is also allowed) cows should be grazing outdoors. The regulations differ within regions of Sweden owing to differences in climate-in the north, for example, dairy cows shall be allowed to graze outside for 60 days, whereas in the middle of Sweden it is 90 days, and in the south 120 days. The definition of a "day" is that the dairy cows come outside and have access to the pasture for at least six hours in a row, during the day or at night.

[1] https://www.government.se/information-material/2020/03/animal-welfare-act-20181192/
[2] https://www.government.se/information-material/2020/03/animal-welfare-ordinance-201966/

[3] Statens jordbruksverks föreskrifter och allmänna råd om nötkreatursh̊allning inom lantbruket m.m.; Statens Jordbruksverks föreskrifter (sharepoint.com)

Assessing the evidence base

Any change to guidance or legislation in respect of livestock farming is best discussed with as complete an evidence base as possible so that there is full transparency of the relative advantages and disadvantages of different production systems. The current review therefore follows best practice guidelines of Cochrane and the Collaboration for Environmental Evidence (Figure 1).

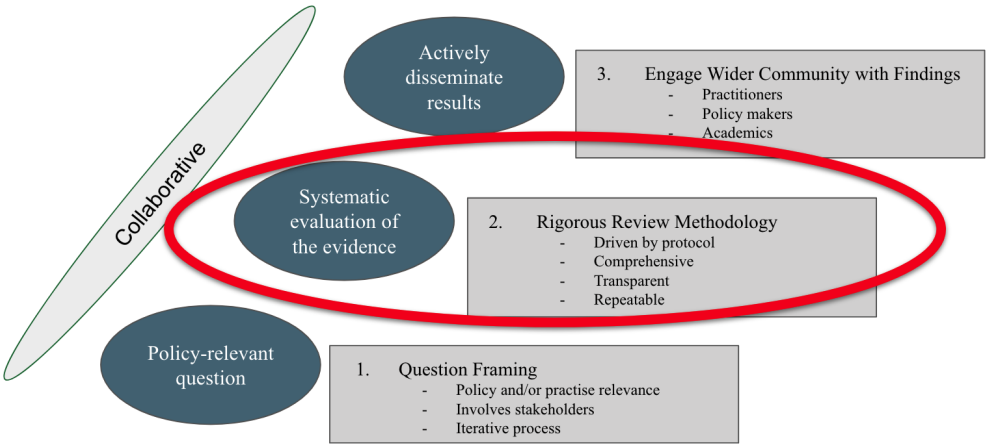

Figure 1: Framework for a systematic evidence evaluation adapted from Collaboration for Environmental Evidence "Guidelines and Standards for Evidence Synthesis…" (CEE, 2022).

Stakeholder engagement and Advisory Group

Following good practice guidelines, the review involved stakeholders and advisors. An initial meeting organised by the Swedish Farmers' Foundation for Agricultural Research outlined the review objectives and solicited interest for participation amongst those attending the meeting. In addition to people who indicated interest from this meeting, potential stakeholders were found by emailing a list of people found through:

1. Authors of relevant academic papers;
2. Contact persons from relevant farmer organisations and groups
3. Contact persons from relevant veterinary or dairy businesses or collectives
4. Contacts suggested by people from any of the groups 1-3, above

A total of 54 people were invited to select times when they were available, if interested, in the first stakeholder meeting during the week commencing 24th January, 2022. The email contained a brief description of the review aims and an Agenda.

The first stakeholder meeting was held on Friday 28th January between 10:00-12:00 CET and attended by 14 people from a range of different organisations engaged in academia, policy-making, practice, and advocacy. The group discussed the focus of the review, developing a search strategy (suggesting keywords and phrases, sources of information, and compiling an exemplar set of publications to verify the search), what literature to include and exclude (including geographical extent and languages of articles to consider). The meeting also discussed the role of an Advisory Group (AG) formed from participants and others who had expressed interest in advising, but who could not attend the stakeholder meeting. Expressions of interest were collected after the meeting by a Google form and an AG of 14 people from 5 countries was formed.

Inputs from the stakeholder meeting were used to draft the Protocol. The draft was circulated to the AG, amended and finalised. The first AG meeting was held in April 2022, and a second in October 2022. Email was used to communicate with AG members between meetings to seek guidance at different stages of the review.

Stakeholder meeting 1: Friday 28th January 2022. Fourteen attendees.
Advisory Group meeting 1: 25th April 2022. Five attendees.
Advisory Group report: 1st July 2022. Sent to all AG members.
Multiple emails and individual Zoom calls during Spring/Summer 2022.
Advisory Group meeting 2: 14th October 2022. Four attendees. (plus two providing input separately)
Stakeholder engagement meeting 2 Stockholm: 27th October 2022

Objective of the review

Primary question:

Are health and welfare of dairy cows kept in all-year loose housing systems better or worse than those kept on pasture, and does non-pasture outdoor access differ in impacts to indoor-only systems (i.e. without outdoor access)?

Secondary question(s):

(i) Are all aspects of natural behaviour exhibited by dairy cows on pasture fully fulfilled by cows permanently in loose housing systems?
(ii) Are there differences in health, welfare and behaviour of dairy cows housed in different systems of all-year loose housing?

A PICO framework has been used to categorise the different aspects of both the primary and secondary questions (Table 1).

Table 1: Definitions of the question (PICO)

PICO	**Description**
Population	Dairy cows in temperate countries
Intervention	All-year loose housing
Comparator	Combination of seasonal loose housing and grazing; different types of loose housing systems; before-and-after introduction of all-year loose-housing
Outcomes	Health and welfare of dairy cows; natural behaviours enabled

Methods

Search Strategy

Bibliographic databases

The search strategy was designed to maximise the amount of relevant literature found by following best practice guidelines outlined in Livoreil et al. (2017). An iterative approach was applied to identify, improve, and optimise keywords and search terms (Appendix 1). The initial list of terms was collected in an online stakeholder workshop (28th January 2022) resulting in a list of 129 keywords or phrases. These terms were expanded upon through the identification of synonyms and preferred terms from relevant literature provided by stakeholders and the advisory group. This extensive list of keywords and terms was then optimised against a test set of 20 relevant articles (Appendix 1) to ensure the maximum return of relevant literature while reducing the overall quantity of irrelevant literature (Appendix 1). The terms from the optimised search were then combined into a boolean string and used to search 5 online bibliographic databases (Table 2). Searches of the online bibliographic databases and aggregates were conducted in English only. Justification for this common limitation is provided in research by Ramírez-Castañeda (2020) that reports 98% of publications in science are written in English. The databases were chosen because they have high coverage rates of veterinary journals and others with significant veterinary content (Grindlay et al. 2012) and extensive coverage of dairy farming journals. The comprehensiveness of the search was assessed during the optimization process by ensuring that test set articles present in Web of Science and CAB Abstracts were returned by combining key terms and keywords in Table 3. Optimization of these combinations of terms was determined by achieving the lowest return rate while finding all test articles. (Appendix 1).

Grey literature

Grey literature and organisational websites (e.g. institutional reports, government reports, non-governmental studies) were also searched to minimise biases arising from the including only published, academic literature (Table 2). An additional call for grey literature was made to stakeholders, the advisory group, and relevant organisations (Table 2).

Table 2: List of sources of literature to be searched

Source of literature	Resource	Location
Online Bibliographic Databases & Aggregators	Web Of Science (Core collection)	www.webofscience.com/
	CAB Abstracts	https://www.cabi.org/
	Scopus	https://www.scopus.com/
	MEDLINE	https://pubmed.ncbi.nlm.nih.gov/
	PubMed	https://pubmed.gov
	Organic eprints	https://www.orgprints.org/
Research Organisations' Websites	Swedish University of Agricultural Sciences Lund University Universitat Kassel Harper Adams University Queen's University Belfast University College Dublin Norwegian University of Life Sciences Natural Resources Institute Finland Wye CollegeUniversity of Copenhagen University of Copenhagen Aarhus University Wageningen, Netherlands INRA, France	https://www.slu.se/en/ https://www.lunduniversity.lu.se/ https://www.uni-kassel.de/uni/en/ https://www.harper-adams.ac.uk/ https://www.qub.ac.uk/ https://www.ucd.ie/ https://www.nmbu.no/en https://www.luke.fi/en/ https://www.imperial.ac.uk/ https://www.ku.dk/ https://international.au.dk/ https://www.wur.nl/en/wageningen-university.htm https://www.inrae.fr/en
Grey Literature	Swedish Board of Agriculture Swiss Federal Veterinary Office The Swedish Veterinary Association (Sveriges Veterinärförbund - SVF) Hushållningssällskapet (https://hushallningssallskapet.se/) Danish Dairy Board AHDB Dairy (UK) SVA (national veterinary institute) Svensk mjölk European Dairy Farmers Swedish Farmers' Foundation for Agricultural Research MS Iceland Dairies (Mjólkursamsalan) TINE (Norway) LRF Federation of Swedish farmers and dairy Sweden Växa Seges Denmark IDF International Dairy Farmers Association Moorepark Animal & Grassland Research and Innovation Centre, Ireland Teagasc, Ireland	

Table 3: Key words/terms used to search bibliographic databases.

PICO Element	Key Term
Population 1	cow* | cattle
Population 2	dairy
Intervention/Comparator	barn*|concrete floor|cubicle*|dry lot|dry-lot|earth floor|forced ventilation|free-stall|freestall*|hard core floor|house|housing|Indoor*|rubber mats|Shed|slat*|sloped system|Stall*|stanchion*|tether*|tiestall|zero graz*|sheds|
Outcome	anti inflammatory|antibiotic use|body condition|feeding compet*|hoof quality|milk amount|milk qualit* & "milk composition"|milk yield*|social licking|activit*|behavior*|behaviour*|cleanliness|comfort*|discomfort*| diseas*|disorder*|distress*|fertility|health|infect*|lame*|lesions|locomotion|longevity|lunging|mastitis|oestrus|parasit* [also helminth*, nematod* etc.|patho*|perching|pregnancy|rumination|stress*|welfare|wellbeing|well-being

Reference checking

We also searched references cited in relevant literature reviews and in background sections of included papers - sometimes referred to as 'snowballing'- to identify other relevant articles to screen (CEE, 2022).

Article screening and study inclusion criteria

Screening strategy

Following searching in each of the bibliographic databases and aggregators, articles were uploaded into EndNote20[4], subscription reference management software published by Clarivate, and Mendeley[5], published by Elsevier that offers some free storage. Duplicate articles were removed and the resulting combined set of articles uploaded into Rayyan[6], a free natural-language processing tool that employs machine learning for screening articles for systematic evidence evaluation. Articles were screened for eligibility at two stages: (i) title and abstract assessment; and (ii) full-text assessment.

Consistency checking

Articles were single- screened by six screeners. In order to check consistency of screening at title and abstract stage, sets of 100 articles were screened by all screeners and inter-rater agreement was assessed using Cohen's kappa. Differences in screening were discussed amongst the screeners and the process repeated with sets of 100 articles until a satisfactory level of agreement was reached (0.6). At full-text screening, sets of 6 articles were similarly assessed by all screeners until inter-rater agreement was reached.

Inclusion criteria

In order to be included in the review, articles had to include information from all elements of the PICO framework. In the first assessment, screening was based on information in the title and abstracts

[4] https://endnote.com/
[5] https://www.mendeley.com/
[6]https://www.rayyan.ai/

in response to three questions: (i) Are the words in the title or abstract, or both, directly related to or included in the review questions? (ii) Does the title or abstract, or both, describe an original research study, as opposed to a literature review or opinion piece? (iii) Was measurement of an outcome of interest (e.g. disease or behaviour) an objective of the study? Only articles with a positive response to all three questions were eligible to proceed to the second stage - screening at full-text. Studies were included if they related to countries or regions with similar production and/or climatic conditions to those of Sweden: Europe (excluding Mediterranean climatic regions), temperate North America, Australia, and New Zealand. There were no exclusion criteria on the basis of date of research.

Exclusion criteria

Studies were excluded if they related to work in: Africa, Central and South America, Asia and Russia. The intention was to limit the language of the original article to those languages that could be read by the review team and advisory group members willing to provide some translation: English, French, German, Spanish, Swedish, Greek, and Arabic. In practice, all of the academic journal articles were in English and a limited number of non-English articles were retrieved in the grey literature process, most of which did not satisfy inclusion criteria. Grey literature was only assessed where there was at least an English summary or English captions on tables and figures, or where Google Translate provided a clear indication of the methods used in the article.

Risk of Bias Assessment

Full-text studies were evaluated for appropriateness of the study design for the research question, and an assessment of specific criteria related to the study design and the likelihood of internal and external biases, adapted from guidance in Pinchbeck & Archer (2020) and will use checklists adapted from the Joanna Briggs Institute[7]and Arnott et al. (2012). The evaluation was conducted without consideration of the study results to avoid interpretation bias. The following questions were posed for risk of bias assessment:

- Is the sample size for each treatment group sufficiently large? (>5)
- Are missing data accounted for?
- Are the criteria for inclusion in the sample clearly defined - including details of randomisation where appropriate?
- Are the study subjects and the setting described in detail?
- Are all of the intervention/exposure and control treatments described including details of treatment, timing, and duration of application?
- Are the criteria for the measurement of all different interventions and their comparators (termed ‘intervention arms’) described in detail?
- Are strategies to deal with confounding factors stated?
- Is there a clear account of the statistical methods used to compare groups for all outcome(s)?
- Are all raw data available?

The results of these assessments were not used for excluding studies from the evidence map, but are intended to enable filtering of studies at risk of bias from specific analyses.

[7] https://jbi.global/critical-appraisal-tools

Data coding and extraction strategy

Data from the included studies were extracted and summarised in a standardised evidence table. In addition to metadata about the article (authors, title, date of publication, source, abstract) taken directly from the bibliographic databases, and study design details, coded by the review team, information based on the PICO elements were extracted by the review team. In total, 266 data fields were completed ; not all studies included data in all fields. The extracted information was used to assess the effects of all-year loose housing on health and welfare of dairy cows. Geographic location data (latitude/longitude expressed in decimal degrees) were either taken directly from the article or added using Google Maps to look up locations of place names in the article. Articles which provided data for multiple interventions were treated as separate 'studies' and all comparisons between housing systems were captured within each study separately. Consistency amongst coders and data extractors was assessed in the same way as full-text article screening, and differences were resolved by repeated discussion until agreement was reached.

Approach to Missing Data

Where data from articles was unclear or aggregated authors were contacted by email for clarification. Disaggregated data or clarified data provided by authors was recorded and included where possible following the same inclusion criteria. Missing data was not imputed.

Potential effect modifiers/reasons for heterogeneity

Several potential effect modifiers that may have contributed to heterogeneity in the outcome were identified during the full-text screening and recorded for all the studies included in this review. These fell broadly into three categories:

1. Population-source variation: herd size, cattle breed, age of study animals,
2. Intervention-source variation: (i) house bedding type (straw, sawdust, rubber mats, slats), house space allowance (m^2 per cow); (ii) distance to walk to exercise area or pasture (if pastured control); (iii) surface of tracks to exercise or pasture (e.g. bark, gravel, earth, concrete); (iv) number of grazing days; (v) strategy (e.g. night-grazing, always open door, water and shadow at the pasture); (vi) Milking-system, milk yield and characteristics; (vii) other forms of environmental enrichment present (e.g. scratch brushes).
3. Location-source variation: i.e. location of farm.

Meta-analysis

Studies were assessed for suitability for meta-analysis. To be deemed 'suitable', studies had to include data for the number of measurements, mean, and standard error or standard deviation. (The Results section explains that there were too few studies of sufficient comparability to facilitate meta-analyses, so these methods are not described further here).

Data synthesis and presentation

Data for date and source of publication, location of study, and study design (including cow breeds, farm/herd size, etc.) are summarised in a series of tables and figures. Data showing numbers of different housing comparators for (i) year round loose housing with zero grazing vs housing with grazing; and (ii) year round loose housing without outdoor access vs housing with outdoor access are presented in tables. HWB data were assessed for positive, negative or no-difference (i.e. direction of effect) following Beaver et al. (2021) for each housing comparator. The number of independent studies from which the HWB measures were taken was assessed and used to filter results on the basis of strength of the evidence base for subsequent statistical and narrative analyses.

CCA analysis and data visualisation were carried out in R (R. Core Team, 2012) using the packages Vegan (Oksanen et al. 2015). The data were square-root converted in order to standardise the distribution and stabilise the variance (Harvey et al. 2019). A square root transform was chosen because it can be applied to data sets containing zero values. The HWB outcome data were subjected to Detrended Correspondence Analysis (DCA) to determine whether a linear or unimodal ordination method should be used (Ter Braak and Prentice, 1988). The CCA analysis was used to cluster the HBW data based on the direction of effect and housing comparator. These groups were used for narrative analyses of the studies.

Online interactive map

Using the geo-locations of all included studies, results are presented on an interactive map. The map is available at https://oxsrev.github.io/evidencemaps/swedish_dairy/. Brief operational details can be found on the website. The map was constructed using Open-Source software (Martin et al., 2017). Each circle on the map represents one location at which an evidence point was generated, or a regional cluster if more than one location occurred nearby. Circle size represents the count of evidence points that occurred at the location. Pie segments represent a percentage of the evidence points at a location for each dairy cow housing type. Users can zoom in on the map to the point where individual circles indicate individual studies. Clicking the circle brings up a summary sheet of the associated studies. Users can also filter the map to show studies on one or a combination of attributes (e.g. type of housing, breed of dairy cows, etc.), and by dimension slices to show numerical filters (e.g. year, number of farms, etc.). Users can also filter the map to show studies on one or a combination of attributes (e.g. type of housing, breed of dairy cows, etc.), and by dimension slices to show numerical filters (e.g. year, number of farms, etc.).

Results

Selection of articles

A total of 166 articles were retrieved and processed according to the methods described above. Figure 2 shows the filtering process at each stage, which resulted in an evidence map of 53 Articles.

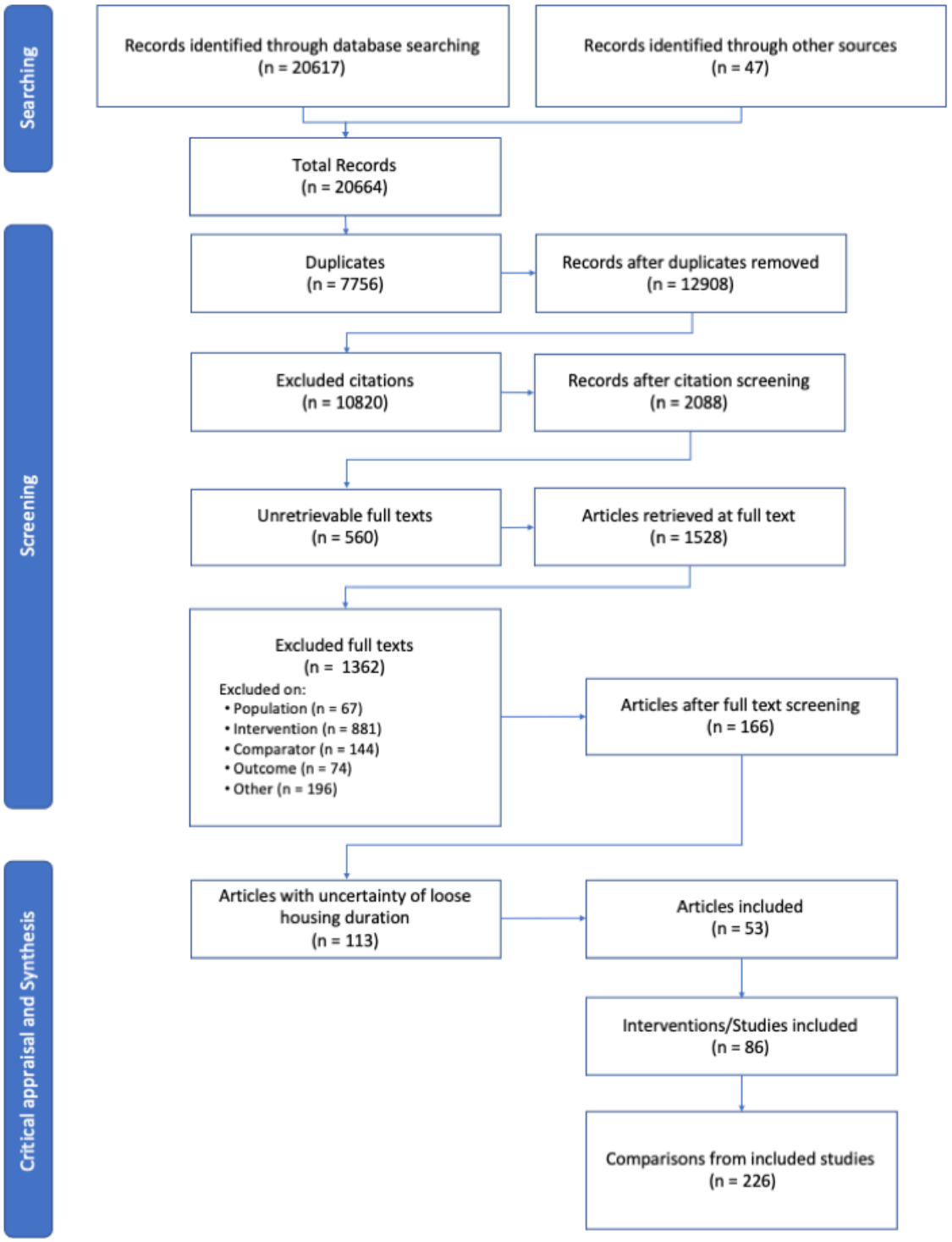

Figure 2: Selection and screening of articles and studies reporting inclusion/exclusion.

Location of studies

Figure 3 shows the location of studies, and the number of studies at the country level.

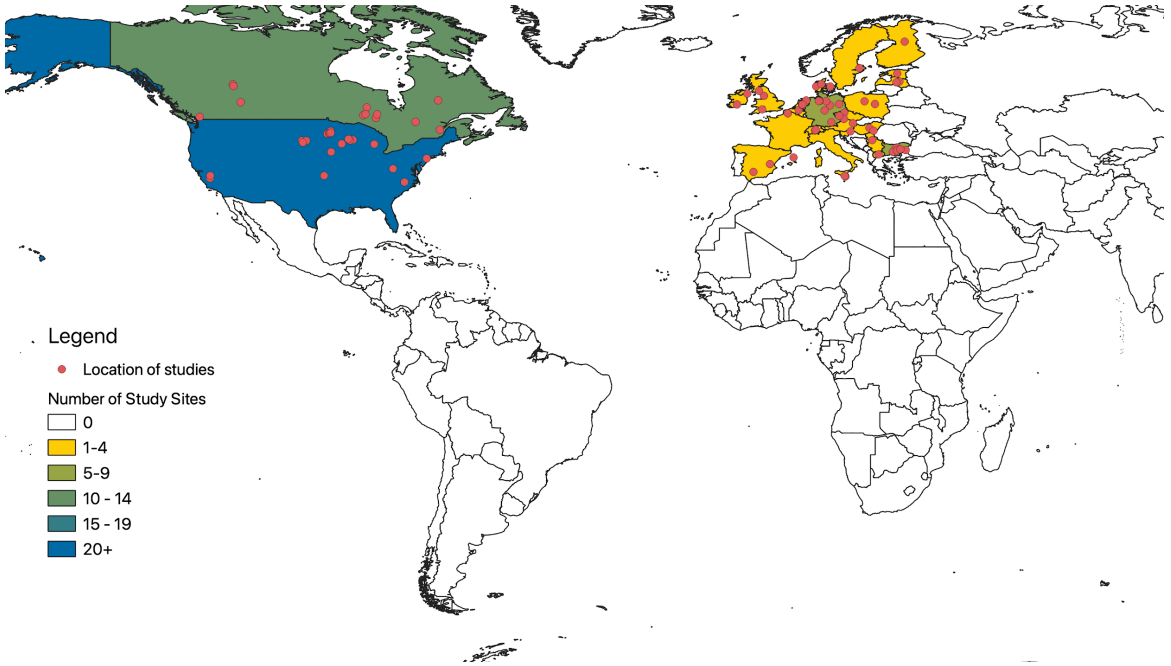

Figure 3: Location of studies included in the evidence base.

Source of studies

Publication type

There are a total of 53 articles of literature that were included at full text of which the vast majority were made up of journal articles (n=52). Data from one conference paper and one book section were also included (Table 4).

Table 4. Number of articles by publication type.

Publication Type	No.
Journal Article	52
Conference Paper/Proceedings	1
Total	53

Key Journals

A total of 27 journals/resources were represented across the included articles of literature. The *Journal of Dairy Science* was the most prominent (n=15), followed by *Animal* (n=4), *Livestock Science* (n=4), and *Preventive Veterinary Medicine* (n=3) (Table 5).

Table 5. Journal/resources with two or more articles.

Journal/Resource Title	**No.**
Journal of Dairy Science	15
Livestock Science	4
Animal	4
Preventive Veterinary Medicine	3
Veterinary Record	2
Veterinary Journal	2
Animals	2
Agricultural Science and Technology	2

Dates of publication

The earliest article included was published in 1997. The years with the highest number of publications include 2012, and 2017 with 6 publications in both years. The years 2006-2017 are best represented across the included articles of literature (Figure 4).

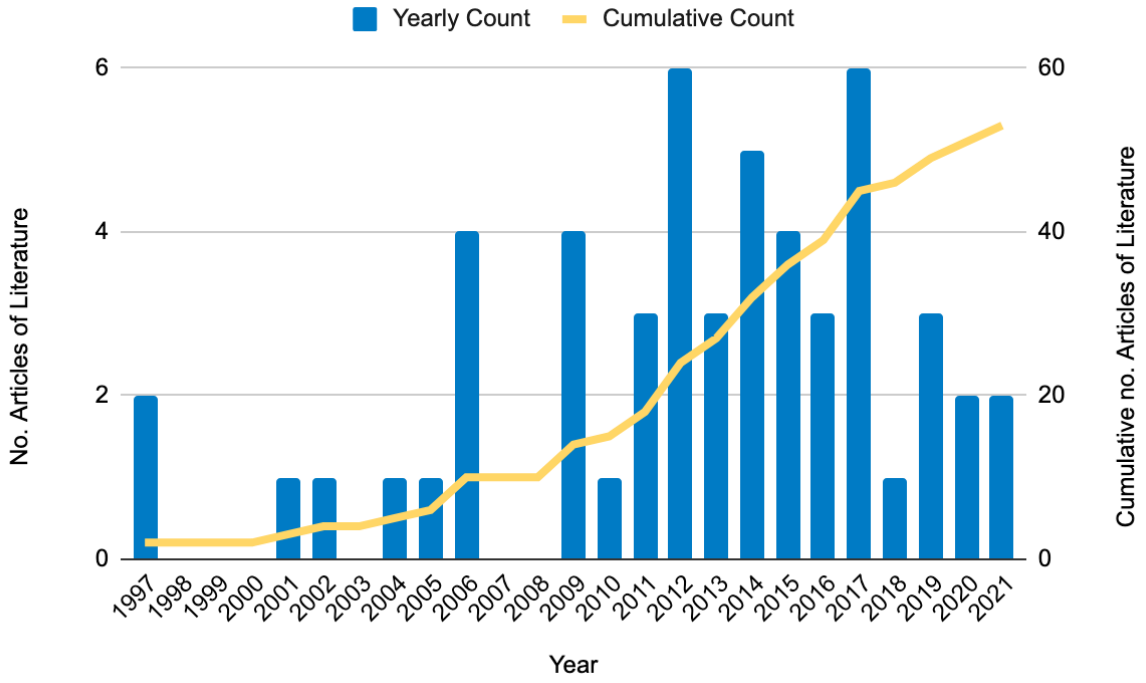

Figure 4: Number of articles of literature by date of publication of articles included in the evidence map by (i) individual year and (ii) cumulatively.

Dairy Cow Breeds

The total number of dairy cow breeds represented were 16 of which the dominant breeds examined in the included articles were Holstein cross (n=30) and Holstein (n=17) cattle (Figure 5).

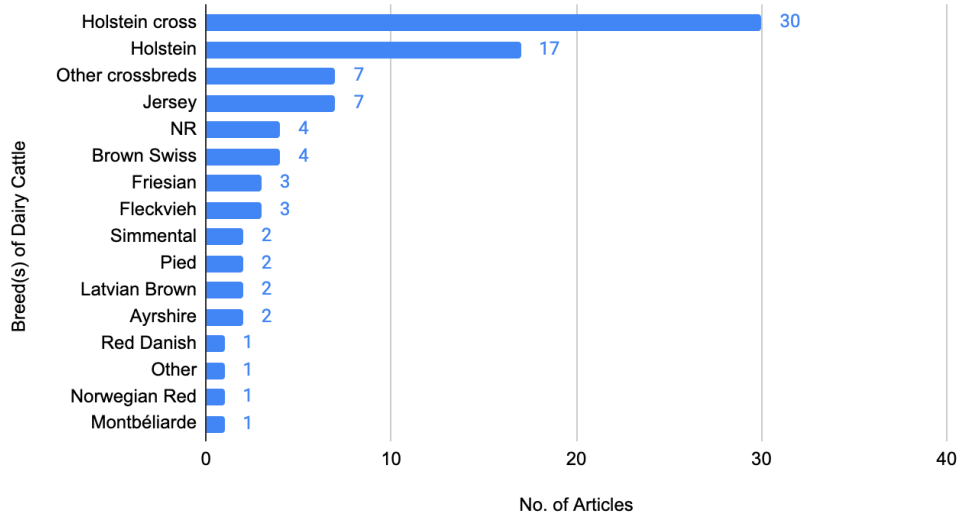

Figure 5: Dairy cow breeds studied in Articles. NR stands for 'not reported', where studies did not specify the breed of dairy cow under consideration.

Risk of bias assessment

Fig 6 shows the scores for studies on the basis of the nine questions posed.

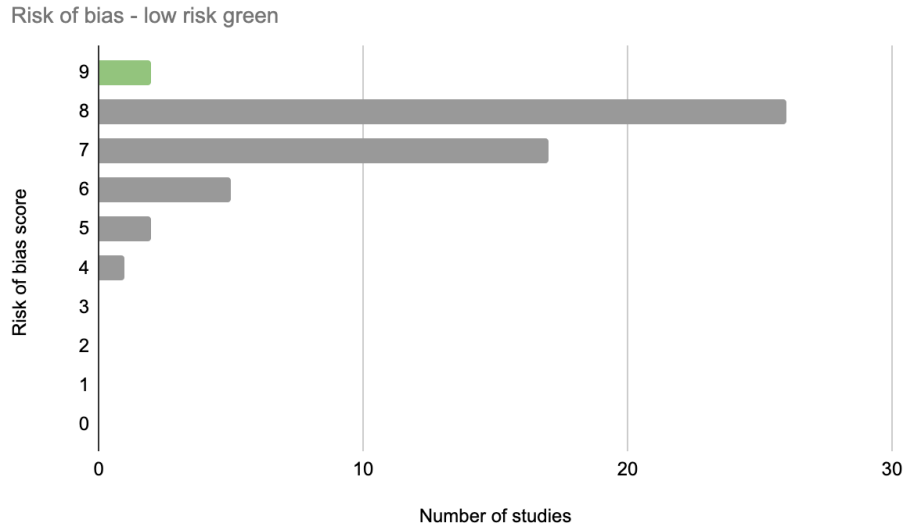

Figure 6: Risk of bias scores for included articles.

Health, welfare, and behaviour measures

From the 53 articles of literature a total of 120 separate HWB outcomes were coded from 86 studies, with 226 comparisons. These include diseases, welfare indicators, assessments of behaviour (including natural behaviours), and risk factors. Across these 120 HWB measures 839 instances of HWB individual measurements were extracted. They fall into 13 broad categories (following Arnott et al. 2017 and Beaver et al. 2021). Some HWB outcomes were derived from zero-grazing vs zero-grazing systems and indoor vs indoor systems which were included in the evidence base and only used to answer the relevant question(s).

Meta-analysis

Meta-analysis was not possible from the collected data due to heterogeneity across collection methods and units for each of the health and welfare outcome measurements (<n=3).

Year round loose housing with zero-grazing vs grazing

Housing comparisons: zero-grazing vs grazing

Following the PICO framework adopted for the review (outlined in Methods), the intervention for all the following analyses was year-round loose housing with zero grazing. The comparators were different housing systems with grazing (Table 6).

Table 6: Number of comparisons of "loose housing with zero grazing" vs "any other housing system (inc loose housing) with grazing".

Number of Comparisons across all included studies	Housing Type - with Grazing/Pasture Access				
	Loose	Tiestall	Mixed	Indoor (Unspecified)	Total
Year Round Loose Housing - Zero Grazing	50	18	2	5	75

Health, welfare, and behaviour measures: zero-grazing vs grazing

Table 7 shows the evidence base for HBW measures (outcomes) for year-round loose housing with zero grazing (HT1) vs housing with grazing (HT2). The data come from 25 articles containing 34 studies, with 75 housing comparisons, reporting 398 health, welfare and behaviour outcomes. The number of incidences are reported as positive (HT1+), negative (HT1-) or no-difference (HT1=HT2). These refer to, respectively: the health measure was better than, worse than, or the same as year-round loose-housing for zero-grazing compared with a housing system with grazing. Caution must be applied, however, in interpreting simple counts of data. Firstly, simple counts of data, based on incidences reported, cannot provide robust evidence that allows for a determination of association (correlation or causation), as outlined in the Method section. Secondly, even allowing for this limitation, a count can be useful to indicate trends within the evidence base, but it also needs to be interpreted with a consideration of the number of independent studies the measurements come from. Table 7, therefore shows not only the raw directional counts, but also shows the number of independent studies that contributed to each HWB count. The colour coding is a visual aid to the relative strength of the evidence, with darker purple indicating higher counts and a gradation of green to red showing high to low numbers of studies. As explained in the Method section, some articles yielded multiple independent studies.

Table 7: The total number of incidences for each health, welfare, and behaviour outcome measure reported in the evidence base (positive, HT1+; negative, HT1-; no-difference, HT1=HT2), and the number of studies these are derived from for year round loose housing with zero-grazing vs housing with grazing.

Category	HWB Measure	HT1+	HT1-	HT1=HT2	Studies n=34
Udder Health	Mastitis	8	10	3	7
	SCC/SCS	12	2	1	5
	Yeast pathogens	3		2	1
	Bacterial pathogens - Streptococcus uberis	4	1		1
	Bacterial pathogens - Streptococcus sp	5			1
	Bacterial pathogens - Streptococcus faecalis	3		2	1
	Bacterial pathogens - Streptococcus dysgalactiae		5		1
	Bacterial pathogens - Serratia species	3		2	1
	Bacterial pathogens - Pseudomonas species	2		3	1
	Bacterial pathogens - Morganella species	1		4	1
	Bacterial pathogens - Klebsiella species	2		3	1
	Bacterial pathogens - Escherichia coli (E. coli)	3	2		1
	Bacterial pathogens - Enterobacter species	1		4	1
	Bacterial pathogens - Corynebacterium species	1	4		1
	Bacterial pathogens - Coagulase +ve staphylococci	5			1
	Bacterial pathogens - Coagulase -ve staphylococci	2	3		1
	Bacterial pathogens - Citrobacter species	1		4	1
	Bacterial pathogens - Bacillus species	2	3		1
	Bacterial pathogens - Arcanobacter pyogenes		5		1
Foot and Leg Disorders	Heel Horn erosion	3	5	2	6
	Digital Dermatitis	1	8	1	5
	White Line disease	1	6	1	4
	Sole haemorrhages		5	3	4
	Sole ulcer			7	3
	Interdigital hyperplasia		5	2	3
	Interdigital Dermatitis	1	3	2	2
	Double Sole		3	3	2
	Severe Hock lesions	2		2	4
	Non-infectious claw disease	3	1		2

Category	HWB Measure	HT1+	HT1-	HT1=HT2	Studies n=34
Foot and Leg Disorders (cont.)	Infectious claw disease	2	2		2
	Leg Diseases	2	1		1
	White line lesion			2	2
	Sole Lesion	2			2
	Maximum axial wall deviation			2	2
	Hock lesions (Unspecified)	2			1
	White Line haemorrhages		1		1
	Swelling		1		1
	Foot Infection			1	1
Uterine Health	Vaginal discharge			6	1
	Dystocia			6	1
	Reproductive tract diseases	3			1
	Retained foetal membranes			1	1
	Resumption of oestrus cyclicity			1	1
	Neospora caninum		1		1
	Metritis		1		1
	Endometritis		1		1
	Chlamydophila abortus			1	1
Maintenance	Dirty Legs/Flanks/Udders		1	8	4
	Hygiene Score	1		3	4
	Dirty legs		1	1	2
	Lying time		1		1
	Dirty udders			1	1
	Cows lying outside cubicles			1	1
	Collisions with cubicles		1		1
Nutrition	Body condition	1		7	3
	Absence of hunger			6	1
	βHB		1		1
	plasma concentrations of triglycerides	1			1
	NEFA		1		1
	Cholesterol	1			1
Respiratory Health	Nasal discharge		3	3	1
	Hampered respiration			6	1
	Coughs			6	1

Category	HWB Measure	HT1+	HT1-	HT1=HT2	Studies n=34
Respiratory Health (cont.)	Bovine herpesvirus specific antibodies (IBR-gE)		1		1
	Bovine herpesvirus specific antibodies (IBR-gB)		1		1
Gut and Intestinal Health	Diarrhoea		1	5	1
	Digestive tract diseases	2	1		1
	Rectal temperature		1		1
	Johne's disease		1		1
	Bovine viral diarrhoea (BVD)		1		1
Behaviour and Stress	Social Behaviour (Unspecified)			6	1
	Human–animal–relationship (Unspecified)			6	1
	Emotional State (Unspecified)			6	1
	Avoidance Distance		4	2	1
	Aggression			6	1
Lameness	Lameness (Unspecified)	2	6	4	7
	Severe Lameness	2	4	5	6
	Locomotion Difficulties		3	4	6
	Clinical lameness		2		2
Skin Health	Hairless Patches		4	2	1
	Lesions and swellings		1	3	4
Ocular Health	Ocular discharge	2		4	1
Mortality	Mortality	2	10	6	5
Metabolic Health	Metabolic disorders (Unspecified)	2	1		1

Canonical Correspondence Analysis: zero-grazing vs grazing

Some HWB outcomes came from single studies, while the data for other HWB outcomes came from multiple studies. The data for the CCA were therefore filtered to include only HWB metrics drawn from more than one study (see Appendix 1 for all studies and all HBW metrics).

A data set of 22 articles containing 33 studies, and 59 comparisons, reporting 192 health and welfare outcomes across 4 housing systems was included in the CCA which explained 3.5% of the data set across CCA1 (1.9%) and CCA2 (1.6%) (Figure 7). CCA1 (horizontal axis) broadly represents positive (left), negative (middle), and no-difference (right) in HBA outcomes. This is due to how the groupings fall; CCA axes are not to be treated as representing a continuum. CCA2 (vertical axis) broadly indicates the split between tiestall and indoor (unspecified) (top), and loose housing (bottom). The detailed statistical interpretation of the CCA is described in the numbered points immediately below, followed by a summarised interpretation of these findings in the subsequent paragraphs.

1. **Somatic Cell Count (SCC), Somatic Cell Score (SCS)**, **sole lesion**, and **non-infectious claw disease** are more closely aligned to CCA1 and are associated with more positive impacts for year round loose housing with zero-grazing vs housing with grazing. **Sole lesion**, and **non-infectious claw disease** outcomes are associated with the comparator loose housing with grazing; and **SCC/SCS** with loose-housing with grazing, and tiestall with grazing.
2. **Heel horn erosion**, **digital dermatitis**, **clinical lameness** and **white line disease** are aligned between CCA1 and CCA2 and are associated with more negative impacts for year round loose housing with zero-grazing vs housing with grazing, and are most associated with the housing comparator loose housing with grazing.
3. **Locomotion difficulties**, **sole ulcer**, **maximum axial wall deviation**, **white line lesion**, **dirty legs/flanks/udders**, **body condition**, **hygiene score** and **lesions and swellings** are more closely aligned to CCA1 and are associated with no-difference impacts for year round loose housing with zero-grazing vs housing with grazing and by the housing comparators loose-housing or mixed (loose and tiestall housing) with grazing.
4. **Mastitis**, **severe hock lesions**, and **infectious claw disease** are aligned between CCA1 and CCA2 and are associated with a combination of positive and negative impacts for year round loose housing with zero-grazing vs housing with grazing. **Mastitis** is most explained by the housing comparator tiestall with grazing, while **infectious claw disease** and **severe hock lesions** are both explained by the housing comparator loose housing with grazing.
5. **Severe lameness**, **dirty legs**, **double sole**, **interdigital dermatitis**, **interdigital hyperplasia**, **sole haemorrhages**, and **lameness (unspecified)** are aligned between CCA1 and CCA2 and are associated with a combination of positive, negative, and no-difference impacts for year round loose housing with zero-grazing vs housing with grazing. **Mortality** is more aligned with CCA2 and is also associated with a combination of positive and negative and no-difference impacts for year round loose housing with zero-grazing vs housing with grazing. The impacts of **mortality** are most associated with the housing comparator tietall housing with grazing, while **severe lameness**, **dirty legs**, **double sole**, **interdigital dermatitis**, **interdigital hyperplasia**, **sole haemorrhages**, and **lameness (unspecified)** are most associated with the housing comparator loose housing with grazing.

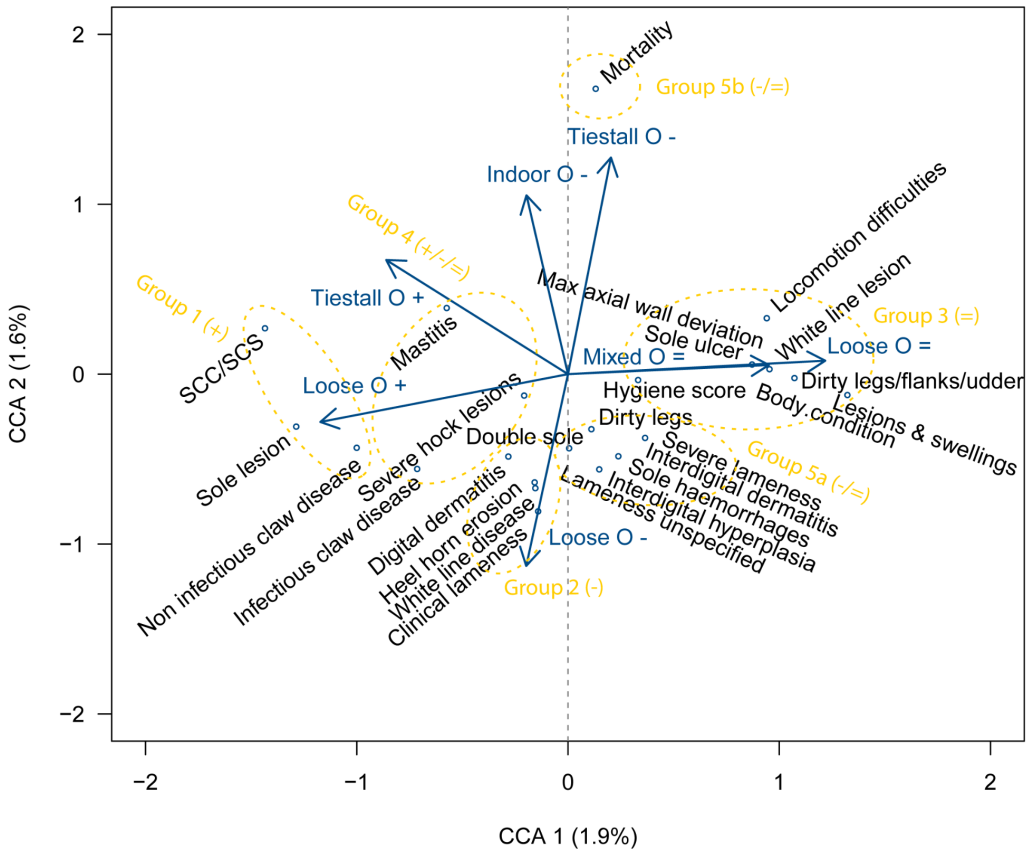

Figure 7: Canonical correspondence analysis of year-round loose-housing with zero-grazing vs housing with grazing and impacts to health and welfare and behaviour of dairy cattle.

<u>Groups: zero-grazing vs grazing</u>

To summarise, five groups can be detected in the CCA (Fig 7) .

Group 1 comprises three HWB outcomes where the evidence base indicates largely positive instances of year-round loose housing with zero grazing compared with housing systems with grazing (though it should be noted that individual studies within this grouping contain instances of the opposite effect or no difference). For **SCC/SCS**, Cielava and Lambertz reported only positive, Bradley reported positive and negative, Pavlenko reported no difference. For **sole lesion,** Baird reported positive. For **non-infectious claw disease**, Haggman reported positive and negative. Cielava and Pavlenko use a tiestall with outdoor access comparator, all others use loose-housing. Bradley, Pavlenko, Baird, Lambertz and Haggman contained a grazing comparator of 91- 180 days duration of grazing, Cielava used 181-270 days and Haggman also used 271-365 days (Table 8).

Table 8: Group 1 health, welfare, and behaviour outcomes: zero-grazing vs grazing.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
1505_Baird_2009	UK	Loose housing	Sole lesion	+
1519_Haggman_2015	Finland	Loose housing	Non-infectious claw disease	+|-
2507_Cielava_2017	Latvia	Tiestall	SCC/SCS	+
294_Bradley_2001	UK	Loose housing	SCC/SCS	+| -
517_Pavlenko_2018	Estonia	Tiestall	SCC/SCS	=
738_Lambertz_2014	Germany	Loose housing	SCC/SCS	+

Group 2 comprises four HWB outcomes where the evidence base indicates largely negative instances of year-round loose housing with zero grazing compared with housing systems with grazing (though it should be noted that individual studies within this grouping contain instances of the reverse or no difference). For **heel horn erosion**, Baird, Olmos and Holzhauer report negative, Armbrecht reports positive, negative and no difference. For **digital dermatitis**, Olmos, Somers and Holzhauer report negative, Armbrecht reports negative and no difference. For **white line disease**, Olmos and Holzhauer report negative, Armbrecht reports positive, negative and no difference. For **clinical lameness**, Otten reports negative. *N.b.* All articles are for loose-housing systems only. Armbrecht, Baird and Holzhauer used a grazing comparator of 91- 180 days duration of grazing, Olmos and Otten used 181-270 days (Table 9).

Table 9: Group 2 health, welfare, and behaviour outcomes: zero-grazing vs grazing.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
1163_Armbrecht_2017	Germany	Loose-housing	Heel horn erosion	+|-|=
			Digital dermatitis	-|=
			White line disease	+|-|=
2074_Olmos_2009	Ireland	Loose-housing	Heel horn erosion	-
			Digital dermatitis	-
			White line disease	-
2539_Holzhauer_2012	Netherlands	Loose-housing	Heel horn erosion	-
			Digital dermatitis	-
			White line disease	-
3320_Somers_2005	Netherlands	Loose-housing	Digital dermatitis	-
1505_Baird_2009	UK	Loose-housing	Heel horn erosion	-
281_Otten_2013	Denmark	Loose-housing	Clinical lameness	-

Group 3 comprises eight HWB outcomes where the evidence base indicates largely no-difference in instances of year-round loose housing with zero grazing compared with housing systems with grazing (though it should be noted that individual studies within this grouping contain instances of the reverse or no difference). For **sole ulcer**, Armbrecht and Holzhauer report no difference. For **maximum axial wall deviation** and **white line lesion**, Baird reports no difference, For **locomotion difficulties**, Olmos and Herlin report negative, Baird and Meul report no difference. For **body condition**, Meul reports no difference, Armbrecht reports positive and no difference. For **dirty legs/flanks/udders**, Gieseke and Meul report no difference, Armbrecht reports negative and no difference. For **lesions and swellings**, Gieseke and Meul report no difference, de Vries reports negative. For **hygiene score**, Salfer reports positive and no-difference. *N.b.* Herlin used tiestall as a comparator, Meul used a mix of tiestall and loose-housing as comparators, other authors used loose-housing, including one study by Herlin. Herlin and Meul used a grazing comparator of 91- 180 days duration of grazing, Armbrecht, Baird, Gieseke, de Vries, Holzhauer used 91-180 days, Olmos and Salfer used 181-270 days (Table 10).

Table 10: Group 3 health, welfare, and behaviour outcomes: zero-grazing vs grazing.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
2061_Meul_2012	Belgium	Tiestall|Loose-housing	Locomotion difficulties	=
			Body condition	=
			Dirty legs, flanks, udders	=
			Lesions and swellings	=
439_Armbrecht_2019	Germany	Loose-housing	Body condition	+|=
			Dirty legs, flanks, udders	-|=
1521_Gieseke_2020	Germany	Loose-housing	Dirty legs, flanks, udders	=
			Lesions and swellings	=
2084_deVries_2015	Netherlands	Loose-housing	Lesions and swelling	-
1505_Baird_2009	UK	Loose housing	Maximum axial wall deviation	=
			White line lesion	=
			Locomotion difficulties	=
2074_Olmos_2009	Ireland	Loose housing	Locomotion difficulties	-
2392_Herlin_1997	Sweden	Tiestall|Loose housing	Locomotion difficulties	-
1163_Armbrecht_2017	Germany	Loose housing	Sole ulcer	=
2539_Holzhauer_2012	Netherlands	Loose housing	Sole ulcer	=
2119_Salfer_2018	USA	Loose housing	Hygiene score	+|=

Group 4 comprises three HWB outcomes where the evidence base is a mix of positive, negative and no-difference instances of year-round loose housing with zero grazing compared with housing systems with grazing, but with some instances of no difference. There is some influence in this group of the tiestall vs loose-housing comparator. For **mastitis**, Gieseke and Washburn report negative, Grabowski and Bradley report positive and negative, Olde-Riekerink reports positive and

no-difference, Armbrecht reports positive, negative and no difference. For **infectious claw disease**, Haggman reports positive and negative. For **severe hock lesions**, Salfer reports positive and no-difference. *N.b.* Washburn used indoors (unspecified) as a comparator, Haggman, Grabowski, Olde-Riekerink, Bradley and Ambrecht used loose-housing as a comparator, Grabowski and Olde-Riekerink also used tiestall as a comparator. Grabowski, Bradley, Olde-Riekerink, Haggman and Armbrecht used a grazing comparator of 91- 180 days duration of grazing, Salfer, used 181-270 days and Washburn, used 271-365 days (Table 11).

Table 11: Group 4 health, welfare, and behaviour outcomes: zero-grazing vs grazing.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
1519_Haggman_2015	Finland	Loose-housing	Infectious claw disease	+|-
2283_Grabowski_1997	Poland	Tiestall	Mastitis	+|-
1521_Gieseke_2020	Germany	Loose-housing	Mastitis	-
2214_Olde-Riekerink_2008	Canada	Tiestall	Mastitis	+|=
294_Bradley_2001	UK	Loose-housing	Mastitis	+|-
3245_Washburn_2002	USA	Indoor (unspecified)	Mastitis	-
439_Armbrecht_2019	Germany	Loose-housing	Mastitis	+|-|=
2119_Salfer_2018	USA	Loose-housing	Severe hock lesions	+|=

Group 5a comprises seven HWB outcomes where the evidence base is mainly negative and no-difference instances of year-round loose housing with zero grazing compared with housing systems with grazing, but with some instances of positive. There is some influence in this group of the tiestall vs loose-housing comparator. For **sole haemorrhages**, Olmos and Holzhauer report negative, Armbrecht reports negative and no difference. For **interdigital hyperplasia**, Armbrecht reports negative and no difference, Holzhauer reports no difference. For **interdigital dermatitis**, Armbrecht reports positive, negative and no difference. For **double sole**, Armbrecht reports negative and no difference. For **lameness (unspecified)**, de Vries and Haskell report negative, Salfer reports positive, negative and no difference, Armbrecht reports negative and no difference. For **severe lameness**, Gieseke reports positive, Salfer reports positive , negative, and no difference, Armbrecht reports negative and no difference. For **dirty legs**, Gieseke reports no difference, Nielsen reports negative. *N.b.* Grabowski and Cielava used tiestall as a comparator, Washburn used indoor (unspecified), and others used loose-housing as a comparator. Nielsen, Haskell, Cielava, Salfer and Olmos used a grazing comparator of 181-270 days duration of grazing, Washburn used 271-365 days, all others used 91-180 days (Table 12).

Table 12: Groups 5a health, welfare, and behaviour outcomes: zero-grazing vs grazing.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
1521_Gieseke_2020	Germany	Loose-housing	Severe lameness	+
			Dirty legs	=
2145_Nielsen_2011	Denmark	Loose-housing	Dirty Legs	-
1163_Armbrecht_2017	Germany	Loose-housing	Sole haemorrhages	-|=
			Interdigital hyperplasia	-|=
			Interdigital dermatitis	+|-|=
			Double sole	-|=
2539_Holzhauer_2012	Netherlands	Loose-housing	Sole haemorrhages	-|=
			Interdigital hyperplasia	=
2084_deVries_2015	Netherlands	Loose-housing	Lameness (unspecified)	-
2108_Haskell_2006	UK	Loose-housing	Lameness (unspecified)	-
2119_Salfer_2018	USA	Loose-housing	Lameness (unspecified)	+|-|=
			Severe lameness	+|-|=
439_Armbrecht_2019	Germany	Loose-housing	Lameness (unspecified)	-|=
			Severe lameness	-|=

Group 5b comprises a single HWB measure, for which the evidence base is mainly negative or no difference, but in contrast to Group 4a, there is a tiestall effect. For **mortality**, Cielava and Washburn report negative, Armbrecht reports no difference, Grabowski reports positive and negative (Table 13).

Table 13: Group 5b health, welfare, and behaviour outcomes: zero-grazing vs grazing.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
439_Armbrecht_2019	Germany	Loose-housing	Mortality	=
3245_Washburn_2002	USA	Indoor (unspecified)	Mortality	-
2074_Olmos_2009	Ireland	Loose-housing	Sole haemorrhages	-

In summary only one HWB measurement (**sole lesion**) showed only positive and no-difference results (and the positives outweigh the no-difference) for zero grazed year round loose-housing. Whereas a total of three indicators (**interdigital hyperplasia**, **sole haemorrhages**, and **clinical lameness**) showed only negative and no-difference results (and the negatives outweigh the no-difference) for zero grazed year round loose-housing.

Year round loose housing without outdoor access vs housing systems with outdoor access (grazing or yards, runs, etc.)

Housing comparisons: without outdoor access vs outdoor access

Housing systems were compared against each other with instances of year round loose housing without outdoor access vs housing with some outdoor access (Table 14).

Table 14: Number of comparisons of loose housing system without outdoor access vs housing system (inc loose housing) with outdoor access (grazing or yards).

Number of Comparisons across all included studies	Housing Type - with Outdoor Access			
	Loose	Tiestall	Indoor (Unspecified)	Total
Year Round Loose Housing Without Outdoor Access	51	33	1	85

Health, welfare, and behaviour measures: without outdoor access vs with some outdoor access

Table 15 shows the evidence base for health, welfare and behaviour measures (outcomes) for year-round loose housing without outdoor access vs housing with outdoor access. The data come from 17 articles containing 21 studies, with 85 comparisons, reporting 332 health, welfare and behaviour outcomes. The number of incidences are reported as positive (HT1+), negative (HT1-) or no-difference (HT1=HT2). These mean, respectively: the health measure was better than, worse than, or the same as year-round loose-housing for zero-grazing compared with a housing system with grazing. Caution must be applied, however, in interpreting simple counts of data. Firstly, simple counts of data, based on incidences reported, cannot provide robust evidence that allows for a determination of causation, or even correlation, as outlined in the Method section. Secondly, even

allowing for this limitation, a count can be useful to indicate trends within the evidence base, but it also needs to be interpreted with a consideration of the number of independent studies the measurements come from. Table 15, therefore shows not only the raw directional counts, but also shows the number of independent studies that contributed to each HWB count. The colour coding is a visual aid to the relative strength of the evidence, with darker purple indicating higher counts and a gradation of green to red showing high to low numbers of studies. As explained in the Method section, some articles yielded multiple independent studies.

Table 15: The total number of incidences for each health, welfare, and behaviour outcome measure reported in the evidence base, and the number of studies these are derived from for year round loose housing without outdoor access vs housing with outdoor access.

Category	HWB Measure	HT1+	HT1-	HT1=HT2	Studies n=21
Udder Health	SCC/SCS	12	2	1	5
	Mastitis	6	2	0	2
	Yeast pathogens	3	0	2	1
	Bacterial pathogens - Streptococcus uberis	4	1	0	1
	Bacterial pathogens - Streptococcus sp	5	0	0	1
	Bacterial pathogens - Streptococcus faecalis	3	0	2	1
	Bacterial pathogens - Streptococcus dysgalactiae	0	5	0	1
	Bacterial pathogens - Serratia species	3	0	2	1
	Bacterial pathogens - Pseudomonas species	2	0	3	1
	Bacterial pathogens - Morganella species	1	0	4	1
	Bacterial pathogens - Klebsiella species	2	0	3	1
	Bacterial pathogens - Escherichia coli (E. coli)	3	2	0	1
	Bacterial pathogens - Enterobacter species	1	0	4	1
	Bacterial pathogens - Corynebacterium species	1	4	0	1
	Bacterial pathogens - Coagulase +ve staphylococci	5	0	0	1
	Bacterial pathogens - Coagulase -ve staphylococci	2	3	0	1
	Bacterial pathogens - Citrobacter species	1	0	4	1
	Bacterial pathogens - Bacillus species	2	3	0	1
	Bacterial pathogens - Arcanobacter pyogenes	0	5	0	1
Foot and Leg Disorders	Heel Horn erosion	3	5	2	6
	Digital Dermatitis	1	8	1	5
	Sole ulcer	0	1	8	4
	White Line disease	1	6	1	4

Category	HWB Measure	HT1+	HT1-	HT1=HT2	Studies n=21
Foot and Leg Disorders (cont.)	Sole haemorrhages	0	5	3	4
	Interdigital hyperplasia	0	5	2	3
	Interdigital Dermatitis	1	3	2	2
	Double Sole	0	3	3	2
	Non-infectious claw disease	3	0	0	1
	Leg Diseases	2	1	0	1
	Infectious claw disease	2	1	0	1
	White line lesion	0	0	2	2
	Sole Lesion	2	0	0	2
	Maximum axial wall deviation	0	0	2	2
	Interdigital Fibroma	0	1	1	1
	White Line haemorrhages	0	1	0	1
	Foot Infection	0	0	1	1
Uterine Health	Reproductive tract diseases	3	0	0	1
	Retained foetal membranes	0	0	1	1
	Resumption of oestrus cyclicity	0	0	1	1
	Neospora caninum	0	1	0	1
	Metritis	0	1	0	1
	Endometritis	0	1	0	1
	Chlamydophila abortus	0	0	1	1
Nutrition	βHB	0	1	0	1
	plasma concentrations of triglycerides	1	0	0	1
	NEFA	0	1	0	1
	Cholesterol	1	0	0	1
Lameness	Locomotion	0	33	1	3
	Locomotion Difficulties	0	1	2	3
	Lameness (Unspecified)	0	1	0	1
	Clinical lameness	0	1	0	1
Gut and Intestinal Health	Digestive tract diseases	2	1	0	1
	Rectal temperature	0	1	0	1
	Johne's disease	0	1	0	1
	Bovine viral diarrhoea (BVD)	0	1	0	1

Category	HWB Measure	HT1+	HT1-	HT1=HT2	Studies n=21
Respiratory Health	Bovine herpesvirus specific antibodies (IBR-gE)	0	1	0	1
	Bovine herpesvirus specific antibodies (IBR-gB)	0	1	0	1
Maintenance	Stockmanship	2	31	1	3
	Lying time	0	1	0	1
Skin Health	Lesions and swellings	0	1	0	1
Mortality	Mortality	2	6	0	2
Metabolic Health	Metabolic disorders (Unspecified)	2	1	0	1
Behaviour and Stress	Social Interaction	5	26	3	3

Canonical Correspondence Analysis: without outdoor access vs outdoor access

Some HWB outcomes came from single studies, while the data for other HWB outcomes came from multiple studies. The data for the CCA were therefore filtered to include only HWB metrics drawn from more than one study (see Appendix 1 for all studies and all HBW metrics). A data set of 13 articles containing 18 studies, with 66 comparisons, reporting 206 health and welfare outcomes across 3 housing systems was included in the CCA which explained 4.3% of the data set across CCA1 (3.3%) and CCA2 (2.8%) (Figure 8). CCA1 (horizontal axis) broadly represents no-difference (left), a mix of mainly negative or no-difference with some positive (middle) and either a more positive or a negative impact (right). CCA2 (vertical axis) broadly indicates a positive impact (top right) or a negative impact (bottom right). The detailed statistical interpretation of the CCA are described in the numbered points immediately below, then summarise the interpretation of these findings in the subsequent paragraphs.

1. ‘Somatic Cell Count (SCC) and Somatic Cell Score (SCS)’, ‘sole lesion’, and ‘mastitis’ are aligned between CCA1 and CCA2 and are associated with more positive impacts for year round loose housing without outdoor access vs housing with outdoor access. ‘Sole lesion’, is most associated with the housing comparator loose housing with outdoor access; and ‘SCC/SCS’ and ‘mastitis’ are associated between the comparators loose-housing and tiestall with outdoor access.
2. ‘Locomotion’, ‘stockmanship’, ‘social interaction’, and ‘mortality’ are aligned between CCA1 and CCA2 and are associated with more negative impacts for year round loose housing without outdoor access vs housing with outdoor access. ‘Social interaction’, and ‘mortality’ are most associated with the housing comparator tiestall with outdoor access, while ‘locomotion’, ‘stockmanship’ are mostly associated with the housing comparator tiestall with outdoor access, and also with some loose housing with outdoor access.
3. ‘Heel horn erosion’, ‘digital dermatitis’, ‘sole ulcer’, ‘white line disease’, ‘sole haemorrhages’, ‘interdigital hyperplasia’, ‘interdigital dermatitis’, ‘double sole’, ‘white line lesion’, ‘maximum axial wall deviation’, and ‘locomotion difficulties’ are more closely aligned to CCA1 and are associated with no-difference or mixed (mainly negative or no-difference with some positive) impacts for year round loose housing without outdoor access vs housing with outdoor access. ‘Heel horn erosion’, ‘digital dermatitis’, ‘sole ulcer’, ‘white line disease’, ‘sole haemorrhages’, ‘interdigital hyperplasia’, ‘interdigital dermatitis’, ‘double sole’, ‘white line lesion’, ‘maximum axial wall deviation’, and ‘locomotion difficulties’ are mostly associated with the housing comparator loose housing with some loose housing with outdoor access.

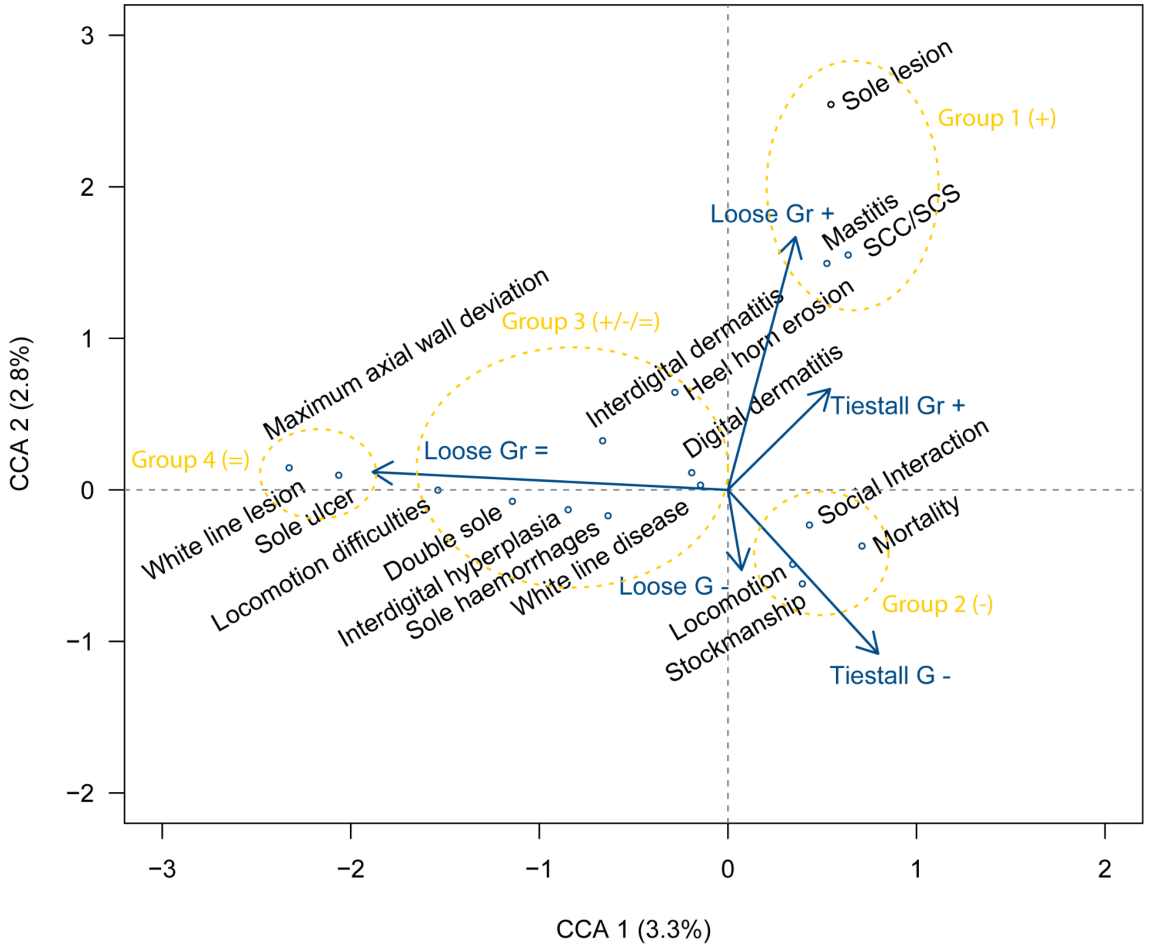

Figure 8: Canonical correspondence analysis of year-round loose-housing without-outdoor access vs housing with outdoor access and impacts to health and welfare and behaviour of dairy cattle.

Based on the CCA (Figure 8) three groups emerge from the data from the evidence base.
Group 1 comprises three HWB outcomes where the evidence base indicates largely positive instances of year-round loose housing with zero outdoor access compared with housing systems with some outdoor access (though it should be noted that individual studies within this grouping contain instances of the reverse or no difference). These derive from six articles. For SCC/SCS, Lambertz and Cielava report only positive outcomes; Bradley reports a mix of positive and negative, and Pavlenko reports no difference. For mastitis, Bradley reports only positive, and Grabowski reports positive if a tiestall comparator is used, but negative for a loose-housing comparator. For sole lesion, Baird reports positive. *Nb.* Cielava and Pavlenko use a tiestall with outdoor access comparator, whereas Bradley, Baird and Lambertz use a loose housing with outdoor access comparator. Grabowski uses both in separate studies. All articles used 91-270 days duration of outdoor access as a component of the comparator (Table 16).

Table 16: Group 1 health, welfare, and behaviour outcomes: outdoor access vs housing with outdoor access.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
1505_Baird_2009	UK	Loose-housing	Sole lesion	+
2283_Grabowski_1997	Poland	Tiestall	Mastitis	+
		Loose-housing	Mastitis	-
2507_Cielava_2017	Latvia	Tiestall	SCC/SCS	+
294_Bradley_2001	UK	Loose-housing	Mastitis	+
			Scc/SCS	+|-|=
517_Pavlenko_2018	Estonia	Tiestall	SCC/SCS	=
738_Lambertz_2014	Germany	Loose-housing	SCC/SCS	+

Group 2 comprises four HWB outcomes where the evidence base indicates largely negative instances of year-round loose housing with zero outdoor access compared with housing systems with some outdoor access. Again, it should be noted that individual studies within this grouping contain instances of the reverse or no difference). These derive from four articles. **Locomotion** and **social interaction** are reported only as negative outcomes (Ornik, Hristov); **stockmanship** is either negative (Ornik), or no difference (Hristov). **Mortality** is negative (Cielava) or negative (Grabowski if in loose-housing), or positive if the comparator is tiestall. *Nb.* Cielava also uses a tiestall comparator, whereas Hristov and Ornik use a loose-housing with outdoor access comparator. Two articles (Ornik and Hristov) contain an outdoor comparator of 271-365 days outdoors. One article uses 91-180 days as the outdoor comparator (Grabowski), one article (Cielava) uses 181-270 days as a component of the comparator (Table 17).

Table 17: Group 2 health, welfare, and behaviour outcomes: outdoor access vs housing with outdoor access.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
2283_Grabowski_1997	Poland	Tiestall	Mortality	+
		Loose-housing	Mortality	-
2507_Cielava_2017	Latvia	Loose-housing	Mortality	-
3201_Hristov_2014	Serbia	Loose-housing	Locomotion	-
			Social interaction	-
			Stockmanship	=
843_Ornik_2010	Slovenia	Loose-housing	Locomotion	-
			Social interaction	-
			Stockmanship	-

Group 3 comprises eight HWB outcomes where the evidence base indicates a mix of positive, negative and no difference for year-round loose housing with zero outdoor access compared with housing systems with some outdoor access. Again, it should be noted that individual studies within this grouping contain instances of positive or negative. These derive from six articles. For **heel horn erosion**, Baird, Olmos and Holzhauer report negative, Armbrecht reports positive, negative and no difference. For **digital dermatitis**, Olmos and Holzhauer report negative, Somers reports negative and positive, and Armbrecht reports negative and no difference. For **interdigital dermatitis** Armbrecht reports positive, negative and no difference. For **white line disease,** Olmos and Holzhauer report negative, Armbrecht reports positive, negative and no difference. For **sole haemorrhages**, Olmos and Holzhauer report negative, Armbrecht reports negative and no difference. For **interdigital hyperplasia**, Holzhauer reports no difference, Armbrecht reports negative and no difference. For **double sole**, Armbrecht reports negative and no difference. For **locomotion difficulties**, Baird reports no difference, Olmos reports negative. One article (Cramer) contains an outdoor comparator of 271-365 days outdoors. Four articles use 91-180 days as the outdoor comparator (Armbrecht, Baird, Cramer and Holzhauer), one article (Olmos) uses 181-270 days and one article (Somers) that uses 1-90 days duration of outdoor access as a component of the comparator (Table 18).

Table 18: Group 3 health, welfare, and behaviour outcomes: outdoor access vs housing with outdoor access.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
1163_Armbrecht_2017	Germany	Loose housing	Digital dermatitis	-|=
			Heel horn erosion	+|-|=
			Interdigital dermatitis	+|-|=
			Sole haemorrhages	-
			White line disease	+|-|=
2074_Olmos_2009	Ireland	Loose housing	Heel horn erosion	-
			Digital dermatitis	-
2539_Holzhauer_2012	Netherlands	Loose housing	Heel horn erosion	-
			Digital dermatitis	-
3320_Somers_2005	Netherlands	Loose housing	Digital dermatitis	-|+
1505_Baird_2009	UK	Loose housing	Heel horn erosion	-

Group 4 comprises three HWB outcomes, where the evidence base shows largely no difference for year-round loose housing with zero outdoor access compared with housing systems with some outdoor access. Again, it should be noted that individual studies within this grouping contain instances of positive or negative. For **white line lesion** and **maximum axial wall deviation**, Baird reports no difference. For **sole ulcer**, Armbrecht and Holzhauer report no difference (Table 19).

Table 19: Group 4 health, welfare, and behaviour outcomes: outdoor access vs housing with outdoor access.

UID First Author Publication Date	Country of study	Housing Comparators	HWB Outcome Measurement	Effect Direction +/-/=
1163_Armbrecht	Germany	Loose housing	Sole ulcer	=
2539_Holzhauer_2012	Netherlands	Loose housing	Sole ulcer	=
1505_Baird_2009	UK	Loose housing	White line lesion	=
			Maximum axial wall deviation	=

In summary only one HWB measurement (**sole lesion**) showed only positive and no-difference results (and the positives outweigh the no-difference) for year round loose-housing without outdoor access. Whereas a total of three indicators (**interdigital hyperplasia**, **sole haemorrhages**, and **locomotion**) showed only negative and no-difference results (and the negatives outweigh the no-difference) for year round loose-housing without outdoor access.

Housing design factors

The evidence base contains even less data that allows for comparisons between different design features of loose-housing systems. Therefore meaningful comparisons can not be made (see Appendix 1 for all the data extracted for this component of the review). Further, given the lack of evidence overall for the benefits or disadvantages of loose-housing over housing systems with grazing or even some outdoor access, it is meaningless in the light of the overarching review question to analyse the data for design features.

Limitations of the Evidence Base

All literature reviews have limitations, and while systematic reviews aim to minimise biases in study selection and synthesis, reviewer bias can also affect the interpretation of results. The first limitation is often the review's scope, which is established by the research topic and the consensus on its components. By including stakeholders in developing the research topic and agreeing a pre-defined methodology (the Protocol) for the review, the systematic method partially overcomes this problem. The Protocol also helped reduce bias in selection of articles for consideration. Terms and phrases were suggested collaboratively, and the sources for published articles were agreed. The extensive search for grey literature helped to some extent to reduce bias caused by citing only academic publications (where a known bias is for positive results to be prioritised over negative or null results). However, it is recognised that the grey literature came largely from organisations known to stakeholders and was thus potentially a limitation.

In common with many other systematic evidence syntheses, despite the large number of articles assessed at full-text for inclusion in the current review, a substantial proportion lacked details of the intervention (year-round loose-housing with zero grazing or no outdoor access), comparator, or disaggregated HWB measures, and were therefore excluded from the review. It is likely that some or many of these excluded articles could be of relevance to the review question, and it is to be hoped that future updates may be able to utilise more disaggregated data.

The lack of comparable data across different studies to allow for meta-analysis is a serious limitation in that no robust statistical analysis of correlation can be made between year-round loose-housing and health, welfare or natural behaviours. Therefore, it is not possible to establish an association between

year round loose-housing on any health, welfare or behaviour measure. Searching for more studies will not change this: there needs to be more consistency in data measuring and data presentation in the primary research. Raw data should be made available to the scientific community to facilitate secondary synthesis.

Discussion

Are health and welfare of dairy cows kept in all-year loose housing systems better or worse than those kept on pasture or with outdoor access?

Assessing the evidence base systematically as a whole, there is no strong evidence of the benefits of keeping dairy cows in year-round loose housing systems as well as no clear evidence of disadvantages of keeping dairy cows in year-round loose housing systems compared with dairy cows with access to pasture and/or outdoor access. There is too much variation across studies in terms of what is measured and how it is measured to allow robust statistical assessment of the direct effects of loose-housing systems and/or grazing on the health, and welfare of dairy cows. Individual studies report detrimental effects for some HWB outcomes, some report beneficial effects for some HWB outcomes, others report no difference between year-round loose housing compared with other systems with outdoor access and/or grazing. It is also important to note that there is no systematic evidence of no-difference in HWB outcomes as a whole between year-round loose-housing and housing systems that allow outdoor access and/or grazing. However, policy and practice changes are best made on the basis of an evaluation of all available evidence, and not individual studies where there is no consensus in the evidence base.

Only one HWB measurement (**sole lesion**) showed a positive trend (meaning the majority of results were positive and there were no negative results) in year-round loose housing with zero grazing and year-round loose housing systems without outdoor access compared with housing systems with outdoor access and/or grazing. Whereas two indicators (**interdigital hyperplasia** and **sole haemorrhages**) showed negative trends (meaning the majority of results were negative and there were no positive results) in both year-round loose housing with zero grazing and year-round loose housing systems without outdoor access compared with housing systems with outdoor access or grazing. An additional negative trend for **clinical lameness** was found for year-round loose housing systems with zero grazing only, and **locomotion** for year-round loose housing systems without outdoor access. However, it must be noted that sole haemorrhages can be seen as a form of sole lesion but opposing trends have been found for these two outcomes, indicating the extent to which data available is limited.

The evidence base does not therefore fully support the overarching findings from Alvåsen's 2015 review, which reported overall positive effects on HWB of giving cows access to pasture during the summer, nor does the evidence base fully align with the review by Arnott et al. (2017) which concluded that cows on pasture-based systems had lower levels of lameness, hoof pathologies, hock lesions, mastitis, uterine disease and mortality compared with cows on continuously housed systems. It is noted however, that these reviews had different study objectives and methodologies and are not therefore wholly comparable with the current systematic review.

A previous European Food Safety Authority synthesis (EFSA, 2009) assembled a number of studies indicating that dairy cattle kept on pasture are healthier, including reduced incidence of mastitis, lameness, hoof problems, trampled teats, and disease incidence, concluding that zero-grazing would increase the likelihood of these health problems. However, the author group (EFSA, 2009) note that: i) it is difficult to identify the causes of these differences, due to large variation between and within housing and grazing conditions across studies making it challenging to identify the factors responsible

for differing health and welfare outcomes, ii) there are risks to welfare in poor grazing conditions and currently little information on the welfare problems associated with pasture-based systems, iii) there is a possibility that some modifications to housing systems would increase the health and welfare of cows in zero-grazing systems to that equal to or above pasture-based systems. However, the evidence base assembled here has demonstrated that studies are available that show both positive and negative effects of all-year round loose housing compared to pasture, indicating the importance of systematic review methodologies to ensure all available evidence is considered to prevent conclusions being drawn that aren't fully supported by all the available evidence .

Are all aspects of natural behaviour exhibited by dairy cows on pasture fully fulfilled by cows permanently in loose housing systems?

There is even less evidence for behaviour indicators and there were none that showed only positive or only negative for year round loose-housing without outdoor access or zero grazing. This supports the findings of Arnott et al. (2017) that there are knowledge gaps, particularly with respect to behaviour and cow preference for grazing over continuous indoor housing.

Clearly, given that one of the five freedoms is freedom to express natural behaviour and grazing is a natural behaviour, access to pasture gives cows the opportunity to behave naturally (Alvåsen, 2015; Arnott et al. 2017), which is by definition denied to cows in year-round loose housing. The current report found limited data on behaviour and it would be challenging to design an experimental study that compared cows' preferences for pasture or indoor housing within a year-round loose housing system.

The finding of Legrand et al. (2009) found that cows showed a strong preference for access to pasture at night and for access to indoor housing during the day when temperature and humidity increased could not be verified in the current review. There were also no studies that investigated another of the five freedoms- free from fear and distress- which may be expected to differ between indoor cows and those on pasture in regions where there may be predator-livestock interactions.

Are there differences in health, welfare and behaviour of dairy cows housed in different types of all-year loose housing?

There were insufficient studies reporting comparable data on housing designs for loose-housing systems to determine which features conferred better HWB outcomes. The current review therefore supports the conclusion by Emmoth (2020) that more controlled studies are needed comparing different housing systems and how they can be developed in order to improve the welfare of dairy cattle.

Implications for future research, policy, and practice

- More controlled studies are needed of different loose housing systems using health, and welfare measures that can be directly compared.
- There is a need for more primary research comparing natural behaviours of dairy cows in loose housing systems and behaviours on pasture.
- Future research: data collected should be made accessible in a disaggregated form to allow for meaningful further analysis to be conducted, beyond the studies they are collected for.
- The current review provides a framework that can incorporate new comparative primary research.

Declarations

The authors declare that there are no conflicts of interest.

Data Availability

All data presented in this report can be found in Appendix 1.

Acknowledgments

Funding for this project has been provided by the Swedish farmers' foundation for agricultural research. The authors would also like to acknowledge those involved in the stakeholder workshops and advisory group: Mattias Norrby, Helena Elofsson, Suzanne Cewe, Isabelle Veissier, Marina von Keyserlingk, Matthew Jordon, Louise Winblad von Walter, Robert Taylor, Sarah Bolton, Nuria Maria Arribas Vera, Tilli Tansey, Gareth Arnott, Elina Åsbjer, and Birgitta Staaf Larsson.

In addition, we acknowledge the help of the following people in finding literature and clarifying results within published papers: Frida Åkerström, Marina von Keyserlingk, Marcia Endres, Jurgen Jansen, John Webster, Ulf Emanuelson, Karin Persson Waller, Silvana Popescu, Peter Raundal, Jason Lombard, Pascal Savary.

References

Alvåsen, K., 2015. Ekonomiska konsekvenser av krav på bete för mjölkkor. SLU, Uppsala, Sweden.
Arnott, G., Roberts, D., Rooke, J.A., Turner, S.P., Lawrence, A.B. and Rutherford, K.M.D., 2012. Board invited review: The importance of the gestation period for welfare of calves: Maternal stressors and difficult births. *Journal of Animal Science*, *90*(13), pp.5021-5034.

Arnott, G., Ferris, C.P. and O'connell, N.E., 2017. Welfare of dairy cows in continuously housed and pasture-based production systems. *Animal*, *11*(2), pp.261-273.

Beaver, A., Weary, D.M. and von Keyserlingk, M.A., 2021. Invited review: The welfare of dairy cattle housed in tiestalls compared to less-restrictive housing types: A systematic review. *Journal of Dairy Science*, *104*(9), pp.9383-9417.

Collaboration for Environmental Evidence (CEE), 2022. Guidelines and Standards for Evidence synthesis in Environmental Management. Version 5.1 (AS Pullin, GK Frampton, B Livoreil & G Petrokofsky, Eds) www.environmentalevidence.org/information-for-authors. [Accessed 25/10/2022]

Emmoth, L, 2020.Självständigt arbete, Sveriges lantbruksuniversitet, SLU, Uppsala, Sweden.

European Food Safety, 2009. "Scientific report on the effects of farming systems on dairy cow welfare and disease." EFSA Journal 7.7 1143r. doi: 10.2903/j.efsa.2009.1143r

European Parliament, Council of the European Union. Regulation (EU) 2016/429 of the European parliament and of the council of 9 March 2016 on transmissible animal diseases and amending and repealing certain acts in the area of animal health ('Animal Health Law'). Off. J. Eur. Union 2016, 84, 1–208. Available online: https://eur-lex.europa.eu/eli/reg/2016/429/oj [Accessed 20/07/2022].

Farm Animal Welfare Council (FAWC), 2009. *Farm animal welfare in Great Britain: Past, present and future*. Farm Animal Welfare Council.

Fraser, D., 2008. Understanding animal welfare. *Acta Veterinaria Scandinavica*, *50*(1), pp.1-7.

Fraser, D., Duncan, I.J., Edwards, S.A., Grandin, T., Gregory, N.G., Guyonnet, V., Hemsworth, P.H., Huertas, S.M., Huzzey, J.M., Mellor, D.J. and Mench, J.A., 2013. General principles for the welfare of animals in production systems: the underlying science and its application. *The Veterinary Journal*, *198*(1), pp.19-27.

Grindlay, D.J., Brennan, M.L. and Dean, R.S., 2012. Searching the veterinary literature: a comparison of the coverage of veterinary journals by nine bibliographic databases. *Journal of veterinary medical education*, *39*(4), pp.404-412.

Harison, R., 1964. Animal machines: The new factory farming industry. *London: Vincent Stuart Publishers. Horowitz, A., C., & Bekoff, M.(2007). Naturalizing anthropomorphism: behavioral prompts to our humanizing of animals. Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals*, *20*(1), pp.23-35.

Harvey, W.J., Nogué, S., Stansell, N., Petrokofsky, G., Steinman, B. and Willis, K.J., 2019. The Legacy of Pre–Columbian Fire on the Pine–Oak Forests of Upland Guatemala. *Frontiers in Forests and Global Change*, *2*, p.34.

Her Majesty's Stationery Office (HMSO), 1965. Report of the Technical Committee to enquire into the welfare of animals kept under intensive livestock husbandry systems. Chairman: Professor F. W. Rogers Brambell. Cmnd. 2836, December 3 1965.

Knierim, U. and Pajor, E.A., 2018. Regulation, enforcement and incentives. *Animal welfare*, (Ed. 3), pp.349-361.

Lawrence, A.B., Vigors, B. and Sandøe, P., 2019. What is so positive about positive animal welfare?—a critical review of the literature. *Animals*, *9*(10), p.783.

Legrand, A.L., von Keyserlingk, M. A., Weary, D.M., 2009. Preference and usage of pasture versus free-stall housing by lactating dairy cattle. Journal of Dairy Science, 92(8):3651-8. doi: 10.3168/jds.2008-1733.

Livoreil, B., Glanville, J., Haddaway, N.R., Bayliss, H., Bethel, A., de Lachapelle, F.F., Robalino, S., Savilaakso, S., Zhou, W., Petrokofsky, G. and Frampton, G., 2017. Systematic searching for environmental evidence using multiple tools and sources. *Environmental Evidence*, *6*(1), pp.1-14.

Martin, A.C., Jeffers, E.S., Petrokofsky, G., Myers-Smith, I. and Macias-Fauria, M., 2017. Shrub growth and expansion in the Arctic tundra: an assessment of controlling factors using an evidence-based approach. Environmental Research Letters, 12(8), p.085007.

Mench, J.A., 2008. Farm animal welfare in the USA: Farming practices, research, education, regulation, and assurance programs. *Applied Animal Behaviour Science*, *113*(4), pp.298-312.

Mench, J.A., 2019. Animal Welfare—Is Intensification the Problem?. In *The Routledge Handbook of Animal Ethics* (pp. 141-153). Routledge.

Oksanen, J.; Blanchet, F.G.; Kindt, R.; Legendre, P.; Minchin, R.; O'hara, R.B.; Simpson, G.L.; Solymos, P.; Stevens, M.H.H.; Wagner, H.; et al. Package 'vegan'. Available online: https://cran.ism.ac.jp/web/packages/ [Accessed 20/07/2022].

vegan/vegan.pdf (accessed on 25 April 2018).

Pinchbeck, G.L. and Archer, D.C., 2020. How to critically appraise a paper. *Equine Veterinary Education*, *32*(2), pp.104-109.

R Core Team. R: *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2012.

Ramírez-Castañeda, V., 2020. Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: The case of Colombian researchers in biological sciences. *PloS one*, *15*(9), p.e0238372.

Robbins, J.A., Von Keyserlingk, M.A.G., Fraser, D. and Weary, D.M., 2016. Invited review: Farm size and animal welfare. *Journal of Animal Science*, *94*(12), pp.5439-5455.

Smid, A.M.C., de Jong, S., Inberg, P.H., Sinclair, S., von Keyserlingk, M.A., Weary, D.M. and Barkema, H.W., 2022. Western Canadian dairy farmers' perspectives on the provision of outdoor access for dairy cows and on the perceptions of other stakeholders. *Journal of Dairy Science*, *105*(5), pp.4461-4473.

Steinfeld, H., Gerber, P., Wassenaar, T.D., Castel, V., Rosales, M., Rosales, M. and de Haan, C., 2006. *Livestock's long shadow: environmental issues and options*. Food & Agriculture Org.

Ter Braak, C.J.; Prentice, I.C. A Theory of Gradient Analysis. *Adv. Ecol. Res.* 1988, 18, 271–317.

van den Pol-van Dasselaar, A., Hennessy, D. and Isselstein, J., 2020. Grazing of dairy cows in Europe—An in-depth analysis based on the perception of grassland experts. *Sustainability*, *12*(3), p.1098.

Van Vuuren, A.M. and Van Den Pol, A., 2006. Grazing systems and feed supplementation. In *Fresh Herbage for Dairy Cattle: The Key to a Sustainable Food Chain* (No. 18, pp. 85-101). Springer.

Vapnek, J. and Chapman, M.S., 2010. Legislative and regulatory options for animal welfare. *FAO legislative study*, (104). Available at https://ssrn.com/abstract=2898362 [Accessed 20/07/2022]

Whiting, T.L., 2013. Policing farm animal welfare in federated nations: The problem of dual federalism in Canada and the USA. *Animals*, *3*(4), pp.1086-1122.