Animal Health Research Reviews

cambridge.org/ahr

The efficacy of litter management strategies to prevent morbidity and mortality in broiler chickens: a systematic review and network meta-analysis

Jan M. Sargeant[1,2], Michele D. Bergevin[1,2], Katheryn Churchill[1,2], Kaitlyn Dawkins[1,2], Bhumika Deb[1,2], Jennifer Dunn[1,2], Dapeng Hu[3], Catherine M. Logue[4], Shannon Meadows[1,2], Carly Moody[1,2], Anastasia Novy[5], Annette M. O'Connor[6], Mark Reist[1,2], Yuko Sato[6], Chong Wang[3,6] and Charlotte B. Winder[1,2]

[1]Department of Population Medicine, Ontario Veterinary College, University of Guelph, ON, Canada; [2]Centre for Public Health and Zoonoses, University of Guelph, ON, Canada; [3]Department of Statistics, Iowa State University, Ames, USA; [4]Department of Population Health, College of Veterinary Medicine, University of Georgia, Athens, USA; [5]Guelph Poultry Veterinary Services, ON, Canada and [6]Department of Veterinary Diagnostic and Production Animal Medicine, Iowa State University College of Veterinary Medicine, USA

Systematic Review

Cite this article: Sargeant JM *et al* (2019). The efficacy of litter management strategies to prevent morbidity and mortality in broiler chickens: a systematic review and network meta-analysis. *Animal Health Research Reviews* **20**, 247–262. https://doi.org/10.1017/S1466252319000227

Received: 5 June 2019
Revised: 27 November 2019
Accepted: 27 November 2019

Key words:
Broiler poultry; litter management; morbidity; mortality; network meta-analysis; systematic review

Author for correspondence:
Jan M. Sargeant,
Department of Population Medicine,
Ontario Veterinary College,
University of Guelph, ON, Canada.
E-mail: sargeanj@uoguelph.ca

Abstract

A systematic review and network meta-analysis (NMA) were conducted to address the question, 'What is the efficacy of litter management strategies to reduce morbidity, mortality, condemnation at slaughter, or total antibiotic use in broilers?' Eligible studies were clinical trials published in English evaluating the efficacy of litter management in broilers on morbidity, condemnations at slaughter, mortality, or total antibiotic use. Multiple databases and two conference proceedings were searched for relevant literature. After relevance screening and data extraction, there were 50 trials evaluating litter type, 22 trials evaluating litter additives, 10 trials comparing fresh to re-used litter, and six trials evaluating floor type. NMAs were conducted for mortality (61 trials) and for the presence or absence of footpad lesions (15 trials). There were no differences in mortality among the litter types, floor types, or additives. For footpad lesions, peat moss appeared beneficial compared to straw, based on a small number of comparisons. In a pairwise meta-analysis, there was no association between fresh versus used litter on the risk of mortality, although there was considerable heterogeneity among studies ($I^2 = 66\%$). There was poor reporting of key design features in many studies, and analyses rarely accounted for non-independence of observations within flocks.

Introduction

Rationale

Litter management strategies serve a vital role in the poultry industry, with the potential to impact bird health, welfare, and productivity (Dunlop *et al*., 2016). 'Litter' refers to the combination of bedding material and fecal matter that typically covers the floor of commercial broiler operations (Shepherd *et al*., 2017). Two key functions of litter in the context of disease control are to absorb moisture from bird excrement, and to minimize exposure of birds to manure to prevent infections that may result in the need to use antibiotics (Shepherd and Fairchild, 2010; Kaukonen *et al*., 2016). Poor litter management can result in the spread and persistence of diseases of major economic importance within and across flocks, such as coccidiosis, necrotic enteritis, and colibacillosis (Yegani and Korver, 2008). Reduced flock health due to wet litter (above 25–35% moisture) may result in decreased weight gain and lower feed conversion efficiency (de Jong *et al*., 2014) and, in worst cases, increased mortality due to compromised respiratory or systemic health.

Prophylactic or metaphylactic antimicrobials may be used alongside litter management strategies to improve flock health by reducing the rates of infectious diseases (Chen and Jiang, 2014). However, it is imperative that antimicrobial use in human healthcare and in animal agriculture be limited to judicious use to reduce the potential for the development of antimicrobial resistance. In an effort to slow the development of antimicrobial resistance, the World Health Organization (WHO) recommends prudent use of antimicrobials in food-producing animals and in human medicine (WHO, 2015). Improved litter management is one strategy that could reduce the need for antimicrobials to control infectious diseases in the poultry industry. Understanding how litter management impacts morbidity, mortality,

CAMBRIDGE UNIVERSITY PRESS

and antibiotic use in broiler chicken production will help producers to make informed management decisions that could reduce the need for antibiotics, while also maximizing the health, welfare, and productivity of the animals.

Systematic reviews provide a transparent and replicable method for identifying and summarizing research evidence from multiple studies (European Safe Food Authority, 2010; Higgins and Green, 2011; O'Connor and Sargeant, 2014; Sargeant and O'Connor, 2014). When results are available from multiple studies that compare interventions against the same outcome, those results may be combined statistically in a pairwise meta-analysis (Higgins and Green, 2011; O'Connor *et al.*, 2014). A pairwise meta-analysis compares the relative efficacy of two interventions, such as two types of treatments or a treatment and a control (Higgins and Green, 2011). Although meta-analysis is useful for synthesizing pairwise comparisons, decision-making may be better informed by an evaluation of the comparative (relative) efficacy of all of the available treatment or intervention options. Network meta-analysis (NMA), an extension of pairwise meta-analysis, provides such a method for evaluating the comparative efficacy of multiple treatment options (Salanti, 2012). In addition, NMA uses information from all comparisons within an intervention network to produce indirect estimates of the relationships between interventions, thus increasing the amount of information available for efficacy evaluations. This analysis approach can also provide estimates of relative efficacy between interventions in cases where no direct comparisons between those interventions have been published. For instance, suppose that there are studies in the literature comparing the efficacy of wood shavings for litter to newspaper and comparing shavings to peat moss. Even if there are no studies directly comparing newspaper and peat moss, this relationship can be inferred indirectly based on the available information on the comparisons between shavings and newspaper, and shavings and peat moss, given certain assumptions (White *et al.*, 2012). Thus, even in the absence of a complete set of direct comparisons between interventions in the literature, NMA can be a valuable tool to inform management decisions.

Objectives

We used a systematic review and NMA to address the review question: 'What is the efficacy of litter management strategies in reducing morbidity, mortality, condemnation at slaughter, or total antibiotic use in broiler chickens?'. While performance measures, including feed conversion and flock uniformity, also are of key importance to growers and veterinarians, we focused our outcomes on those related to disease and therefore of direct relevance to antibiotic stewardship.

Methods

Protocol and registration

A protocol was prepared *a priori* and was reported in accordance with PRISMA-P guidelines (Moher *et al.*, 2015). The protocol was published to the University of Guelph's institutional repository (https://atrium.lib.uoguelph.ca/xmlui/handle/10214/10046) and also is available through the Systematic Reviews for Animals and Food (SYREAF) website (http://www.syreaf.org/contact/). This review was reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for NMA (PRISMA-NMA) guidelines (Hutton *et al.*, 2015).

Eligibility criteria

Primary research studies, both published and non-published (grey literature), available in English were eligible. In addition, eligible studies must have included the following elements based on the PICOS components:

Population (P): Broiler chickens in commercial or research flocks;
Intervention (I): Litter management strategies (as defined by the authors);
Comparator (C): Alternate litter management strategy, or a no intervention control;
Outcomes (O): Clinical morbidity (as defined by the authors), condemnations at slaughter, mortality, total antibiotic use;
Study design (S): Controlled trials with natural disease exposure and analytical observational studies.

Information sources

The electronic databases searched were AGRICOLA (via ProQuest, 1970 to search date), CAB Abstracts and Global Health (via the University of Guelph CAB interface, 1900 to search date), MEDLINE (via PubMed; 1946 to search date), Conference Proceedings Citation Index – Science (via Web of Science, 1990 to search date), and Science Citation Index (via Web of Science, 1900 to search date). No restrictions were implemented in the search strategy related to publication date, language of publication, or study type. In addition, a single reviewer handsearched all available table of contents of the Western Poultry Disease Conference and Poultry Science Association conference proceedings from 1980 to 2018.

Search

The search strategy initially was developed for the Science Citation Index (Web of Science) and comprised search terms related to three concepts: broilers, litter management, and disease prevention. A list of relevant search terms was compiled for each concept within the search strategy; search terms within each concept were linked using the Boolean operator 'OR', and the concepts were linked using the Boolean operator 'AND'. The full search strategy as applied in the Science Citation Index website is provided in Table 1. Database searches were conducted through the University of Guelph library on 18 June 2018. Search results were uploaded to EndNoteX7 (Clarivate Analytics, Philadelphia, PA, USA) and duplicate records were identified and removed. Records were then uploaded to DistillerSR (Evidence Partners Inc., Ottawa, ON, USA) and again de-duplicated. DistillerSR was used for eligibility screening, data extraction, and risk of bias assessment.

Study selection

Titles and abstracts were screened for eligibility. Reviewers were trained on a pre-test of the first 250 citations to ensure the clarity of understanding and consistency of application of the screening questions. Thereafter, two reviewers independently evaluated each citation, using the following questions to assess eligibility:

(1) Is this a primary study that evaluated litter management to reduce clinical morbidity, condemnations at slaughter, mortality, or antibiotic use in broilers?
Yes, No, Unclear

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

Table 1. Full electronic search strategy used to identify studies of litter management strategies to prevent morbidity and mortality in broilers in Science Citation Index (Web of Science) conducted on 18 June 2018

#1 TS = (Chicken* OR Poultry* OR flock* OR gallus* OR broiler*)	193,862 4
#2 TS = ((cake AND litter) OR (caking AND litter) OR (de-caking AND litter) OR (litter AND (sand OR 'sunflower husks' OR 'rick husks' OR shavings OR straw OR sawdust OR paper OR 'orange peel')) OR 'litter reuse' OR 'reused litter' OR 'wet litter' OR 'litter management' OR 'fresh litter' OR 'reused-litter' OR 'litter condition' OR 'recycled litter' OR 'litter treatment' OR 'litter condition' OR 'litter quality' OR 'bedding' OR 'litter moisture' OR 'litter deterioration' OR 'litter type' OR 'litter quantity' OR 'litter friability' OR 'litter drying' OR 'litter flowability' OR 'litter moisture' OR 'litter score' OR 'litter amendment')	17,674
#3 TS = ((litter OR bedding) AND (enzyme OR 'NSP-degrading' OR xylanase OR 'phytatedegrading' OR phytase OR 'electrolyte balance' OR 'dietary Ca' OR 'dietary Na' OR 'dietary P' OR 'dietary K' or 'dietary potassium' OR 'dietary phosphorus' OR 'dietary calcium' OR 'dietary sodium' OR 'water quality' OR 'hard water' OR betaine OR trimethyl OR chloride OR sulphate OR sulfate OR 'sodium bisulfate' OR acidification OR 'aluminum sulfate' OR clay OR diatomaceous earth OR 'heat drying' OR 'ferric sulphate' OR 'ferric sulfate' OR KLASP OR 'poultry guard' OR 'all clear' OR peat OR windrowing))	23,455
#4 TS = (morbidity OR mortality OR antibiotic OR antimicrobial OR dysbacterosis OR dermatitis OR 'foot-burn' OR pododermatitis OR FPD OR 'paw score' OR 'paw quality' OR 'gait score' OR 'foot pad' OR 'hock burn' OR 'breast blister' OR coccidiosis OR Eimeria OR clostridium OR coli OR Escherichia OR coliform OR colisepticaemia OR colibaccillosis OR coligranuloma OR Hjarre's OR 'air sac disease' OR cellulitis OR peritonitis OR salphingitis OR osteomyelitis OR peritonitis OR salpingitis OR synovitis OR panopthalmitis OR omphalitis OR enteritis OR bronchitis OR 'bursal disease' OR proventriculitis OR runting OR stunting OR 'hemorrhagic septicemia' OR 'respiratory disease' OR 'swollen head syndrome' OR coliform OR osteomyelitis)	2,020,977
#2 OR #3 38,793	
#1 AND (#2 OR #3) 2642	
#1 AND (#2 OR #3) AND #4 664	

(2) Is there a concurrent comparison group? (i.e. controlled trial with natural or deliberate disease exposure or analytical observational study)?
Yes, No, Unclear
(3) Is the full text available in English?
Yes, No, Unclear

Citations were excluded if both reviewers responded 'no' to any of the questions; agreement was at the include or exclude decision level. Disagreements were resolved by consensus. Two independent reviewers conducted the full-text eligibility screening of remaining studies, using the first 10 citations as a pre-test. The full-text screening stage included the initial three questions with only yes or no (exclude) options, and additionally:

(4) Is the full text available with >500 words?
Yes, No
(5) What best describes the intervention?
Feed/nutritional associations with litter quality
Direct litter management (litter or flooring type, litter depth, use of fresh versus re-used litter) (NB: the inclusion of this question to broadly categorize the interventions represents a protocol deviation, which was included due to the volume of literature identified after title and abstract screening).
(6) Are at least one of the following outcomes described: morbidity, condemnations at slaughter, mortality, or antibiotic use?
Yes, No
(7) Eligible study design: what is the study design?
Analytical observational study,
Controlled trial with natural disease exposure,
Controlled trial with deliberate disease induction

Agreement for the full-text eligibility screening was at the question level, with conflicts resolved by consensus or by mediation by JMS or CBW if an agreement could not be reached.

Studies were included in the meta-analysis if sufficient data were reported to enable the calculation of the log odds ratio (OR) and the standard error of the log OR, based on the extraction of the prioritized metrics. Additional criteria for inclusion in the meta-analysis are described in the statistical analysis section.

Data collection process

There were a few protocol deviations at the data collection stage. First, as only three observational studies were identified, and controlled trials are considered a higher level of evidence, data were only collected from controlled trials with natural disease exposure. Second, due to the volume of literature and heterogeneity between the intervention types, we extracted data only from studies evaluating litter management (litter or floor type, litter depth, or used versus fresh litter). Studies which evaluated the indirect effects of feed or nutrition on litter quality were excluded at full-text screening as 'not the intervention of interest'.

Data from citations that were eligible following full-text screening were independently extracted by two reviewers using a standardized form, which was piloted on the first five articles by all reviewers. Discrepancies in data extraction were resolved by consensus, or by mediation by JMS and CBW if an agreement could not be reached.

Data items

Study characteristics

Study-level data that were extracted from each eligible trial included study design, publication year, country, month(s) and year(s) during which data were collected, whether the trial was conducted in a research or commercial flock, strain and sex of birds, number of flocks or farms enrolled, inclusion and exclusion criteria at the flock-level, and rearing conditions (e.g. conventional, organic, antibiotic-free).

Interventions and comparators

Details on the interventions, including a description of the intervention, number of birds/rooms/flocks enrolled, age when

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

intervention was initiated, duration of intervention, and concurrent therapy were recorded. Losses to follow-up for each intervention group were captured when reported in the trial.

Outcomes and results

Data were extracted for clinical morbidity (as defined by the study authors), condemnations at slaughter, mortality, and total antibiotic use. It was anticipated *a priori* that there would be variation in the method of presentation of results, and that some trials would present the results in multiple ways. Therefore, the type of results to be extracted was prioritized such that only one set of results for each outcome were extracted per trial. The first priority for the extraction of results was an adjusted summary effect (adjusted OR or risk ratio (RR)). Adjustment referred to covariates or adjustment for non-independence in grouped housing. Variables included in the adjustment, and the corresponding precision estimates were recorded. If an adjusted measure was not reported, unadjusted summary effect size (second priority) or raw data (third priority) were recorded, with applicable variance components. Data presented without variance measures, and for which a measure of variance could not be calculated, were not extracted. If eligible outcomes were measured at multiple time points, data from the final time point were extracted.

Geometry of the network

We used a visual approach to qualitatively evaluate the geometry of the network, to determine if some pairwise comparisons were more common than others and to determine whether the network appeared to have a star or web-like structure. We also evaluated whether there were intervention comparisons that were not linked to the network (i.e. did not have an intervention in common with one or more other published trials).

Risk of bias in individual studies

The Cochrane risk-of-bias tool for randomized trials (RoB 2.0, 2016 version) was used to assess bias at the outcome level for all outcomes and all trials included in the NMA (Higgins *et al.*, 2016). This tool provides a framework for assessing the likelihood that bias could be introduced in each of the five domains. The domains of bias are: bias arising from the randomization process, bias due to deviation from intended interventions, bias due to missing outcome data, bias in measurement of the outcome, and bias in the selection of the reported result. Signaling questions are a component of the RoB 2.0 tool that is used to elicit information on the use of trial features that are relevant to the potential for bias in a trial. Signaling questions were modified for use in poultry trials. A question in the bias due to randomization domain in the Cochrane tool pertains to the method for generating the random allocation sequence. We modified this question to include a response category for studies where the authors reported that allocation to intervention groups was 'random', but did not provide details on the actual method for generating the random sequence. The RoB 2.0 domain bias related to deviations from the intended intervention has a question on whether the participants were aware of their treatment assignments; this always was answered as 'no', in that the 'participants' in these trials were broiler chickens. Another question in this domain asks whether study personnel were blinded, and for the purposes of this review, this question was clarified to refer to blinding of the animal caregivers.

The overall risk of bias within each domain was calculated based on the algorithms suggested by Higgins *et al.* (2016), with one exception. For bias due to the randomization process, we excluded a consideration of allocation concealment because all animals within a flock would be expected to be included in the type of trial involved in this review. Furthermore, it is unlikely that a producer or investigator would have any treatment preference for a given flock, as the differential economic value of a flock would not be known at the time of allocation. This approach has been used in a previous evaluation of livestock trials (Moura *et al.*, 2019).

Summary measures

The baseline risk used to convert the ORs to the RR was obtained by using the distribution of the placebo group and a Bayesian approach method of estimation. For the model related to footpad lesions, posterior mean and standard deviation of the baseline risk mean were −0.3759 and 3.1808. The posterior mean and standard deviation of the baseline risk standard deviation were 2.9462 and 0.7094. For the model related to mortality, posterior mean and standard deviation of the baseline risk mean were −3.4327 and 0.8097. The posterior mean and standard deviation of the baseline risk standard deviation were 0.788 and 0.0911. When studies had zero cells for some data points, the ORs could not be calculated; thus, the trial results could not be included in the analyses.

Planned method of statistical analysis

After completing data extraction, a treatment map of all reported interventions was compiled for litter type, flooring type, and litter additives (Table 2). Litter types and flooring types that were similar were collapsed into the same intervention for the analyses. For the analyses of litter and flooring type, different depths or new versus used litter for the same litter type were collapsed into the same intervention category.

All interventions related to litter type, flooring type, or litter additives contributed to the NMA. However, due to a large number of interventions and low amount of replication, NMA results are shown for single litter types and for additives that were evaluated in more than one study for mortality and the presence or absence of footpad lesions (the most commonly reported morbidity outcome).

The methodological approach for conducting NMA has been described in detail elsewhere (Dias *et al.*, 2011; O'Connor *et al.*, 2013). As the outcomes were binary, results presented as raw data or ORs were converted to the log OR. If the authors reported an RR, this was converted into a log OR using the reported risk of disease in the placebo group. If the authors reported the probability of an outcome in each intervention group based on a statistical model, that probability was converted back to logs OR, using a process described elsewhere (Hu *et al.*, 2019).

Selection of prior distributions for Bayesian analysis

The choice of prior probability distributions was based on an approach reported previously (Dias *et al.*, 2011). Accordingly, we assessed both $\sigma \sim U$ (0.2) and $\sigma \sim U$ (0.5). The results suggested that $\sigma \sim U$ (0.5) was preferred, and so we retained this prior in the model.

Implementation and output

All posterior samples were generated using Markov Chain Monte Carlo (MCMC) simulation, which was implemented using Just

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

Table 2. Treatment map of litter type, floor type, and litter additive interventions for a systematic review and network meta-analysis of litter management strategies to prevent morbidity and mortality in broilers

Intervention category	Description	Code used in analyses
Litter type		
Paper	Paper, newspaper, pelleted paper	PAPER
Grass	Grass, Bermuda grass, switchgrass, Napier grass	GRASS
Leaves	Chopped palm leaves, leaves, refused tea leaves	LEAF
Straw	Wheat straw, wheat stocks, chopped straw, un-chopped straw, paddy straw, rice straw, soybean straw, rapeseed straw, corn stock straw	STR
Shavings	Wood shavings, pine shavings, sawdust, pine bark, wood shavings and sawdust combined	SHAV
Husks	Rice husk, rice hull, paddy husk, groundnut hulls, sunflower hulls, coconut hulls, rice bran	HUSK
Peat moss		PEAT
Sand		SAND
Corn cobs		COB
Kenaf core		KENAF
Sugarcane		CANE
Coir dust		COIR
Chopped maize		CORN
Silage maize		SIL
Gypsom		GYPS
Combinations	Combinations of 2 or more litter types	SH_HU, HU_PA, ST_SH, SH_LE, SH_SA, SH_GY, ST_HU_SH, ST_SA, CA_SH, CA_HU, SH_GR, KE_SH, SH_PA
Unknown	Litter type not reported	NR
Flooring type		
Metal mesh	Wire mesh, steel mesh	FL_A
Plastic	Plastic mesh, perforated plastic	FL_B
Styrofoam	Perforated Styrofoam	FL_C
Wood	Perforated wood	FL_D
Dowelling	Padded, rigid, or rotating	FL_E, FL_F, FL_G
Other (single study)	Solid wood, concrete, dirt, wood shavings and plastic, slatted floors	FL_H, NO_BED, FL_J, FL_K, FL_L
Litter additive		
Acrochips		AGRO
Aluminum		ALUM
SoftAcid		SA
Acidified clay		AC
Sodium bisulfate		SB
Micropan		MP
Other (single study)	Aluminum plus CaCO2, Cellulose, Canola oil, Litteraid, TEZ, volcanic rock, copper sulfate, potassium, zolite, Litterguard	AL_CA, CELL, OIL, LA, TEZ, CSULF, K, ZO, LG

Another Gibbs Sampler (JAGS) software (Plummer, 2015). All statistical analyses were performed using R software (version 3.5.2) (R Core, 2015). The model was fitted using JAGS; JAGS was called from R via the rjags package (version 4-8) (Plummer, 2015). Three chains were simulated in the analysis, and the convergence was assessed using Gelman–Rubin diagnostics. We discarded 5000 'burn-in' model iterations and based all inferences on a further 10,000 model iterations. The model output included all possible pairwise comparisons using log ORs (for inconsistency assessment), RRs (for comparative efficacy reporting), and the treatment failure rankings (for comparative efficacy reporting).

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

Assessment of model fit

The fit of the model was assessed based on the log OR, by examining the residual deviance between the predicted values from the NMA model and the observed value for each study (Dias *et al.*, 2010).

Assessment of inconsistency

NMA relies on an assumption of consistency between direct and indirect intervention effects, apart from the usual variation that stems from a random-effects meta-analysis model (White *et al.*, 2012). For example, if one trial compares the direct effect of a treatment A with the effect of treatment B, and another study compares the efficacy of treatments B and C, then the effect of A relative to B and B relative to C can be used to infer the (indirect) effect of A relative to C. The assessment of inconsistency compares whether the direct effects and the indirect effects give a similar result. We used the back-calculation method to assess the consistency assumption in our NMA (Dias *et al.*, 2010). We compared the estimates from the direct and indirect models and considered the standard deviation of each estimate, rather than relying on the *P*-values.

Risk of bias across studies (across the network)

A modification of the Grading of Recommendation Assessment, Development and Evaluation (GRADE) approach for NMA was used to describe the quality of the network (Salanti *et al.*, 2014; Papakonstantinou *et al.*, 2018). The GRADE framework provides a method of evaluating the quality or certainty of evidence and the strength of the recommendations derived from that evidence. The GRADE for NMA was conducted using the Confidence in Network Meta-Analysis (CINeMA) online software program (http://cinema.ispm.ch). This platform uses a frequentist approach to calculating intervention effects, which is based on the meta-for package in R (Viechtbauer, 2010). Thus, the contribution matrix of direct and indirect evidence produced for the risk of bias in CINeMA is based on a frequentist analytical approach. The GRADE approach in CINeMA evaluates the evidence network for the following domains: within-study bias, across-studies bias, indirectness, imprecision, heterogeneity, and incoherence. These domains were modified for this review by reporting the contribution of studies based on the reporting of randomization and the reporting of blinding of outcome assessors, rather than the domains of within-study bias and indirectness. The rationale was that randomization and blinding would have more variability among studies in this review compared to the overall risk of bias, and would therefore be more informative. In a GRADE assessment for NMA, indirectness refers to the differences between the populations, interventions, and outcomes in the included studies and the populations, interventions, and outcomes that were the target of the NMA (Salanti *et al.*, 2014). We assumed that the trials included in this review would have minimal indirectness because we restricted the review to relevant populations. To characterize randomization as reported in each of the included studies, we sorted each trial included in the NMA into one of three categories as follows: (1) the authors reported random allocation to intervention groups and provided information on the method used to generate the random sequence, (2) the authors reported random allocation to intervention groups without providing information on how the sequence was generated, or (3) the allocation method was not random or no information on the method of allocation was provided. For blinding, we categorized the trials based on the following: (1) that outcome assessors were blinded, or (2) that outcome assessors were not blinded or no information on blinding was provided. The risk of bias instrument that was used for this review also included a question on whether animal caregivers were aware of the intervention allocation (i.e. were not blinded). However, with litter management, it would be difficult to blind caregivers to the intervention. Thus, we did not include blinding of caregivers as a component of this assessment of blinding.

The process required to assess across-studies bias in an NMA is not well developed. Additionally, no pairwise comparisons in this review included more than 10 trials, which is the number typically believed to be necessary for an accurate across-studies bias assessment (Sterne *et al.*, 2000). Thus, we did not access across-studies bias.

The assessment of imprecision in GRADE indicates whether the boundaries of the confidence intervals for the intervention effect fall within or between estimates that would be consistent with a clinically appreciable benefit or harm, or whether the intervention effects are clinically ambiguous (i.e. the confidence intervals span values representing both benefits and harms, or benefit or harm and a null value). We used an OR of 0.8 to represent a clinically meaningful difference. Thus, an appreciable benefit would correspond to an OR of <0.8 and appreciable harm would correspond to an OR of >1.25. ORs of 0.8 and 1.25 also were used in the assessment of heterogeneity because, for risk of bias assessment in an NMA, the major impact of heterogeneity is whether it will affect decision making. We did not present the results of the incoherence analysis from CINeMA, which measures the consistency of the network, because we presented the consistency analysis results based on the Bayesian analysis described previously in the methods section, rather than the frequentist method used by CINeMA.

Additional analyses

A pairwise meta-analysis was conducted to compare the effect of fresh litter and re-used litter on broiler mortality. The pair-wise meta-analysis was conducted in R. 3.5.2 (R Foundation for Statistical Computing, Vienna, Austria) using RStudio version 1.1.463 (RStudio Inc., Boston, MA, USA) with the 'meta' package (Schwarzer, 2019). The meta-analysis used a random-effects approach, and inverse variance was used for weighting. Heterogeneity between studies was assessed using the I^2 statistic (Higgins and Green, 2011; Schwarzer, 2019).

Results

Study selection

The flow of studies through the review process is shown in Fig. 1. Of the 1582 unique references identified in the search, 343 were assessed for eligibility at the full-text screening. Following the full-text screening, 126 eligible trials were identified in 103 publications, of which 97 trials in 76 publications had extractable data for one or more relevant outcomes; thus 97 trials were ultimately included in the review.

Study characteristics

The study characteristics, category of intervention, and outcome types are presented in Supplementary Table S1. Citation details

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

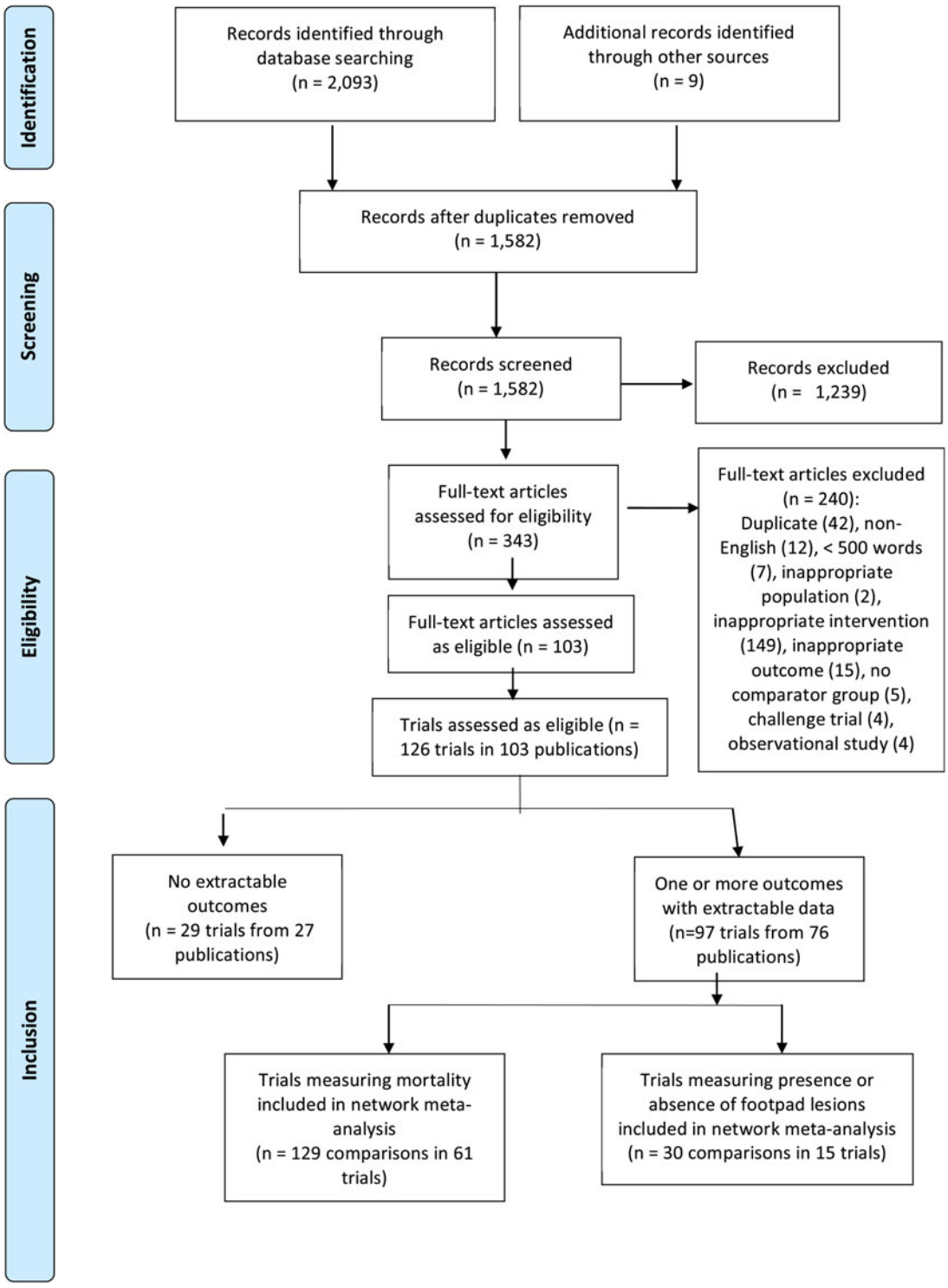

Fig. 1. Flow diagram of included studies for litter management strategies in broiler chickens.

for the studies described in this table are included as a Supplementary reference file. There were 17 countries represented at the trial level, although the country where the trial was conducted was not reported for 60 of the 97 trials. Of the trials where the country was reported, the most common country was the USA ($n = 13$ trials) followed by Brazil, Serbia, and Romania ($n = 3$ trials each). The studies were conducted in university or research flocks in 40 trials, in commercial flocks in 19 trials, and the setting was not reported for 38 trials. The month and year of during which the trial was conducted were not reported for the majority of the trials ($n = 86/97$). The number of farms included in each trial ranged from 1 to 4, with 85 trials conducted on a single farm. In terms of the interventions investigated in each trial, there were 50 trials that evaluated the type of litter, 22 trials that evaluated litter additives, 11 trials that evaluated litter depth, 10 trials that compared fresh to re-used litter (three of which also

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

evaluated litter type or litter depth), six trials that evaluated flooring type, and one trial that evaluated windrowing (i.e. a form of partial composting in which litter is stacked, causing the internal temperature of the stacks to rise and thus destroy harmful pathogens).

Mortality was measured as an outcome in 82 trials (Supplementary Table S1). One or more morbidity outcomes were measured in 54 trials. Morbidity outcomes that were measured on a continuous scale included footpad lesion scores (n = 16 trials), breast lesion scores (n = 7), hock lesion scores (n = 5), gait scores (n = 2), leg lesion scores (n = 1); morbidity lesions measured dichotomously (i.e. presence/absence) included footpad lesions (n = 23 trials), breast blisters (n = 17), bruises (n = 3), leg lesions (n = 2), hock burn (n = 2), scabs (n = 1), and abnormal gaits (n = 1). Condemnations at slaughter were measured in two trials, and there were no trials evaluating total antibiotic use.

Results of individual studies included in the network

Due to a large number of trials and outcomes, the NMAs focused on interventions related to litter type, flooring type, or additives for the two outcomes for which there were the most data: mortality and a morbidity outcome, the presence or absence of footpad lesions. Not all of the trials that measured a given outcome were included in the NMA, because some intervention arms were collapsed into a single category which meant the trial was no longer comparative, and because some trials had zero cells which precluded the calculation of an OR. Of the 82 trials in which mortality was an outcome, 14 trials did not examine litter type, flooring type, or additives as an intervention (Supplementary Table S1), and four trials from three publications had zero cells for all intervention comparisons (Akpobome and Fanguy, 1992; Hafeez *et al.*, 2009; Avdalovic *et al.*, 2017). Additionally, the intervention arms were collapsed into the same intervention based on the treatment map for three of the trials (Do *et al.*, 2005; Abreu *et al.*, 2011; Mendes *et al.*, 2011). Thus, there were 61 trials included in the NMA for mortality as an outcome. None of the included trials included adjustment for flock (barn) effects.

Of the 23 trials that measured footpad lesions as a binary outcome, one trial had all zero cells (i.e. no birds had footpad lesions) (Sahoo *et al.*, 2017) and all birds were positive for the outcome in another trial (Vargas-Galicia *et al.*, 2017); relative effect sizes could not be calculated in either case. There were five trials from four publications in which the intervention arms collapsed into the same intervention based on our treatment map (Petek *et al.*, 2010; Gholap *et al.*, 2012; Avdalovic *et al.*, 2017; Shepherd *et al.*, 2017). One trial had no intervention arms that also were reported in any other trial (i.e. the intervention arms in this trial did not link to the network) (Çavuşoğlu *et al.*, 2018), and therefore this trial could not be included in the comparative analysis. Thus, there were 15 trials from 12 publications included in the NMA for the presence or absence of footpad lesions. None of the trials included an adjustment for flock (barn) effects.

Risk of bias within studies by outcome

A summary of the individual study-level risk of bias for the 61 trials contributing to the mortality outcome NMA is shown in Supplementary Fig. S1. All studies included in that NMA were rated at a high risk of bias due to the method of allocating birds to different intervention or control groups: either the studies did not report the method used to generate the random sequence or the studies did not describe the method of allocation; and the studies also did not report whether there were any baseline imbalances. For three of the risk of bias domains, all of the trials were rated as 'some concerns' about the potential for bias, because in all cases, there was insufficient information reported in the studies to adequately assess the domains. First, none of the trials reported the blinding of outcome assessors, nor did any of the studies provide the additional information needed to assess whether there was the potential for differential management between groups. Further, none of the studies described whether any differences in management were balanced between intervention groups, which is related to the domain of bias due to deviations from intended intervention. Second, insufficient information was reported in each study to evaluate the potential for bias due to missing outcome data. Third, the risk of bias due to selected outcome reporting could not be assessed based on the available information, because assessing this domain requires *a priori* trial protocols to be available, and these are rare in the animal health literature. Another risk of bias domain, which concerns bias due to the outcome measurement process, was considered to be low in all trials in this NMA because mortality is objectively measured.

The results of the risk of bias assessment for the studies included in the NMA for the presence or absence of footpad lesions were similar to the results for the NMA with mortality as an outcome. All except one of the trials included in the NMA were judged to be at a high risk of bias due to the method of intervention allocation, and there were some concerns about the potential for bias due to deviations from intended interventions, missing outcome data, and the selection of the reported result. Unlike the mortality outcome NMA, all studies were rated as a high risk of bias under the domain related to the measurement of the outcome, as morbidity outcomes are more subjectively measured than mortality outcomes and outcome assessors were not blinded to the different intervention groups in any study.

Results of the network meta-analyses: mortality

Geometry of the network

There were 61 trials involving 129 comparisons that reported a mortality outcome and contributed to the NMA (Fig. 1). The geometry of the network is shown in Fig. 2, with the labels for the interventions defined in Table 2. There appeared to be two distinct clusters in the network: one primarily involving litter types that was centered around wood shavings as the type of litter, and the other primarily involving litter additives that was centered around aluminum as an additive. There was one comparison of two flooring types that was not connected to the network, meaning that there was no intervention arm in common with other published studies. The network for all intervention arms for which there were mortality outcomes with non-zero cells is shown in Fig. 3, with the number of comparisons for each intervention shown in parentheses beside the intervention node. Shavings and husks were the most common intervention arms, with 45 and 19 comparison arms, respectively.

Results of individual studies and synthesis of results

Although the NMA was informed by data from all intervention arms related to litter type, flooring type, or litter additives for which mortality was measured as an outcome, relative results for litter management options are presented only for single litter

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

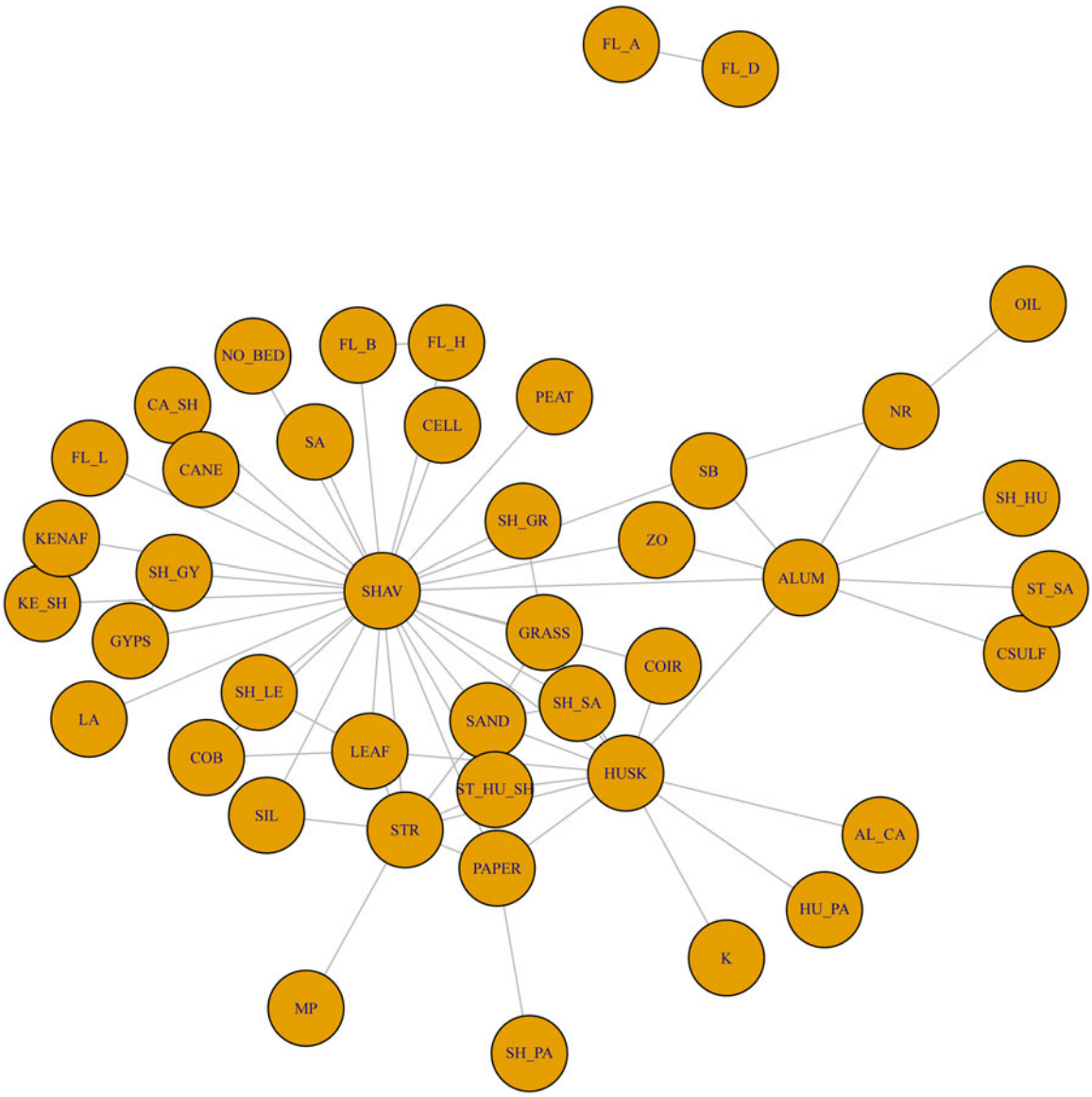

Fig. 2. Geometry of the network of interventions related to litter type, flooring type, or litter additive and their effect on mortality in broiler chickens. Each circle represents an intervention, with a line between two interventions meaning that there was one more comparison between the interventions in the included literature. The key for the intervention acronyms is in Table 2.

types or additives with multiple comparisons (aluminum and sodium bisulfate) due to the large number of possible comparisons. Comparative results are therefore presented for 15 litter types and two additives in the following ways: Fig. 4 illustrates the relative ranking of the litter management strategies for preventing mortality, with the 95% credibility intervals; the RRs for all pairwise comparisons are available in Supplementary Table S2; and the mean rankings are available in Supplementary Table S3.

Based on the wide and overlapping credibility intervals on the relative rankings, there were essentially no differences in mortality risks among the litter management options. Additionally, the imprecision of estimation (indicated by wide credible intervals) for the relative risks means there is insufficient information to determine if there are differences in mortality risks across litter management interventions. These results are consistent with the distribution of the probability of failure (mortality) for each litter management option (see Supplementary Fig. 2).

Exploration of inconsistency

The measures of consistency between the direct and indirect comparisons of interventions with mortality as an outcome are presented in Supplementary Table S4. There was no evidence of inconsistency between the direct and indirect estimates because the credible intervals for all comparisons included the null value. However, as a result of the small number of studies contributing to each estimate, the confidence intervals were wide and the estimates were therefore imprecise.

Results of the network meta-analyses: presence or absence of footpad lesions

Geometry of the network

There were 15 trials with 30 comparisons that reported the presence or absence of footpad lesions as an outcome and contributed to the NMA (Fig. 1). The network for morbidity due to footpad lesions is shown in Fig. 5, with the number of comparisons for

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

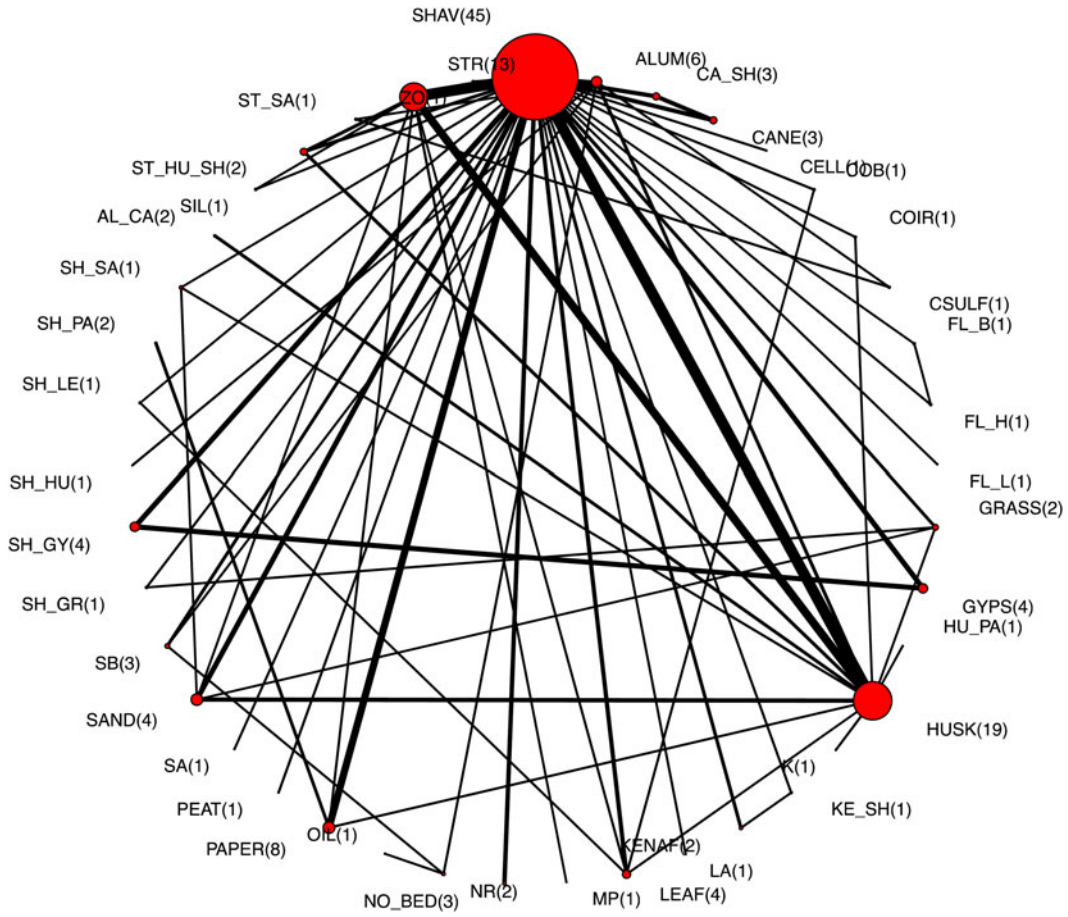

Fig. 3. The network of intervention arms used in a network meta-analysis of the efficacy of litter management strategies to prevent mortality in broiler chickens. The size of the circle provides an indication of the relative number of intervention arms, the width of the lines provides a relative indication of the number of direct comparisons between interventions that were reported in the literature, and the number of arms for each intervention is shown in parentheses beside the intervention node. The key for the intervention acronyms is in Table 2.

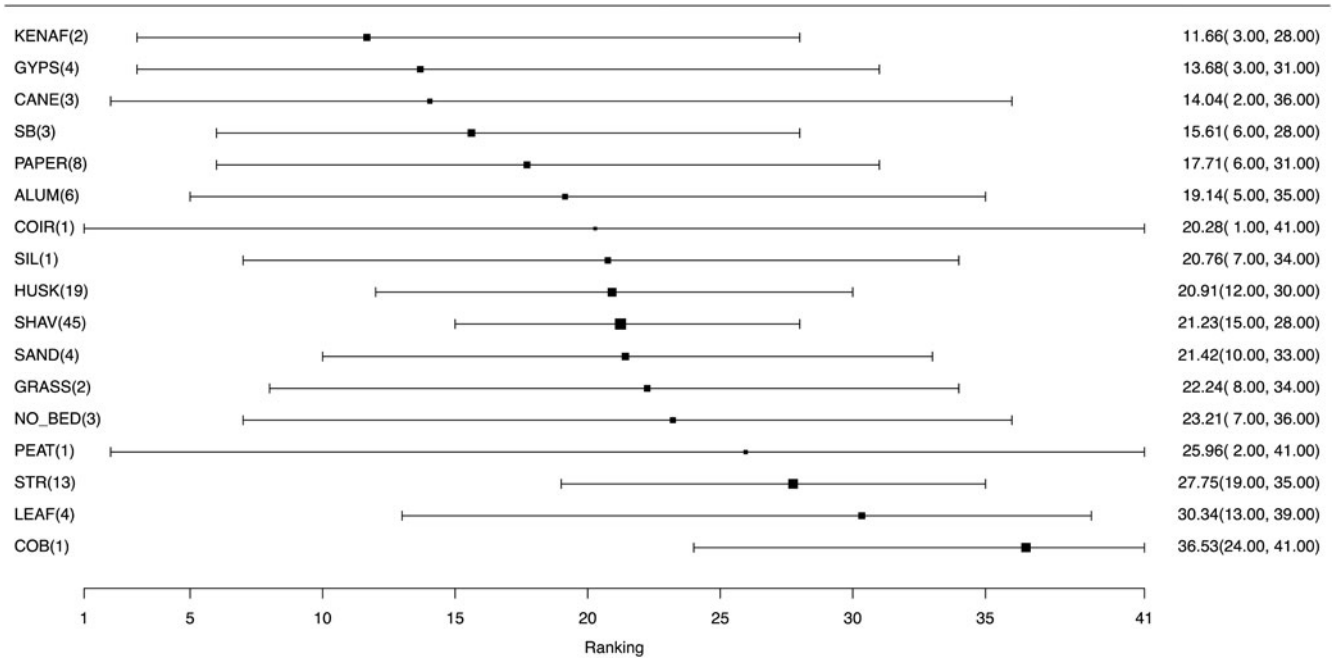

Fig. 4. Forest plot for ranking of intervention arms for litter management options evaluating mortality as an outcome for litter management in broiler chickens. Mean rank and 95% credibility intervals are shown.

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

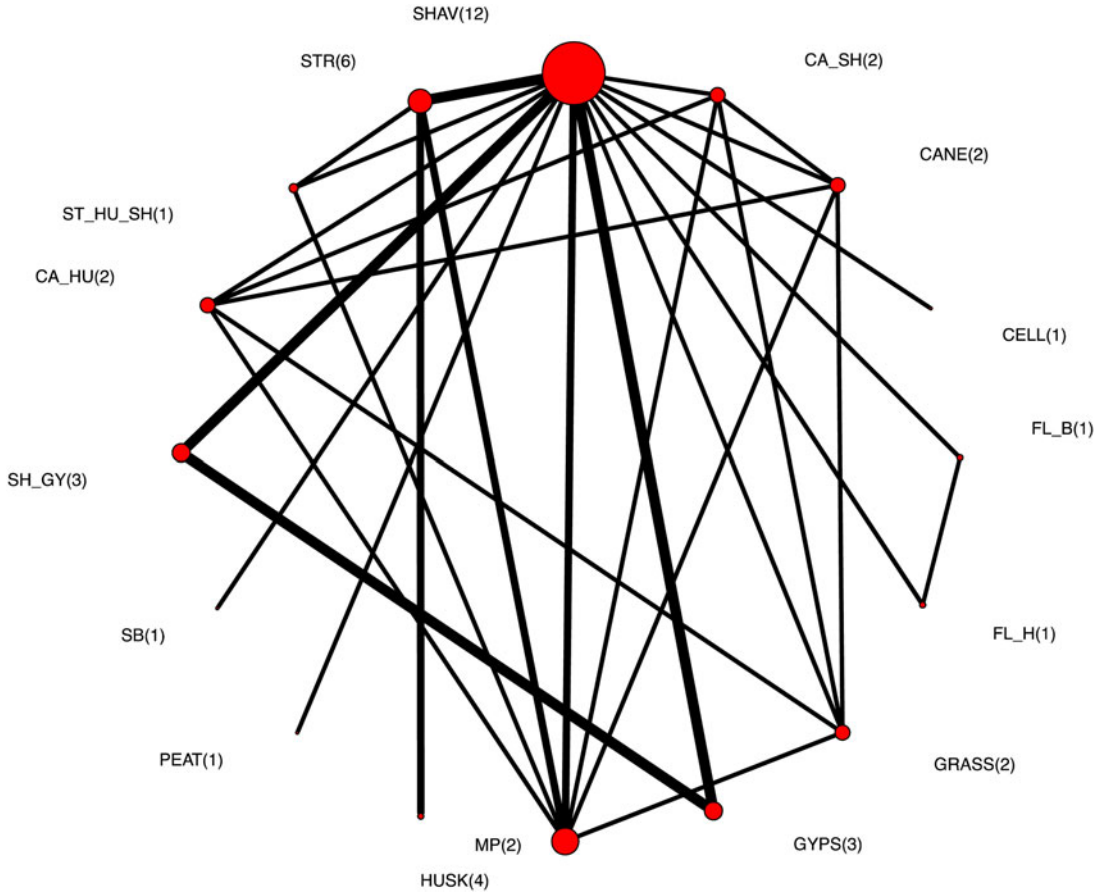

Fig. 5. The network of intervention arms used in a network meta-analysis of the efficacy of litter management strategies to prevent footpad lesions (measured as a dichotomy) in broiler chickens. The size of the circle provides an indication of the relative number of intervention arms, the width of the lines provides a relative indication of the number of direct comparisons between interventions that were reported in the literature, and the number of arms for each intervention is shown in parentheses beside the intervention node. The key for the intervention acronyms is in Table 2.

each intervention shown in parentheses beside the intervention node. The key for the acronyms is in Table 2. Similar to the network for mortality, wood shavings were the most commonly reported intervention, with many of the studies linking to this intervention arm. This suggests a non-random pattern of intervention comparisons in the literature, which is a researcher preference for including wood shavings as a comparator.

Results of individual studies and synthesis of results

Although the NMA was informed by data from all intervention arms related to litter type, flooring type, or litter additives with foot pad lesions measured as a binary outcome, relative results for litter management options are presented only for single bedding types or additives with multiple comparisons (sodium bisulfate) for consistency with the mortality analysis. As a result, comparative results are presented for seven litter types and one additive.

Figure 6 illustrates the relative ranking of each litter management option for preventing footpad lesions, along with the 95% credibility intervals. RRs for the pairwise comparisons of the litter management options are available in Supplementary Table S5; the mean rankings are available in Supplementary Table S6. Peat moss had the highest estimated rank; however, this should be interpreted with caution because this was based on a single comparison. Because of the generally small number of comparisons for any given intervention, the credibility intervals on the rank estimates were wide and overlapping. Thus, there is no compelling evidence that any litter type is superior to the others. These results are consistent with the distribution of the probability of failure (presence of a footpad lesion) for each litter management option (see Supplementary Fig. 3).

Exploration of inconsistency

The consistency measurements between the direct and indirect comparisons for footpad lesions as a binary outcome are presented in Supplementary Table S7. There were a few comparisons where there was evidence of inconsistency between the direct and indirect estimates for several comparisons ($n = 3/30$), because the credible intervals did not include the null value. However, the comparison of the direct and indirect estimates for the remainder of the intervention comparisons indicated consistent results.

Risk of bias across studies

The results of the assessments of risk of bias across studies included in either the NMA for mortality or the NMA for the presence or absence of footpad lesions are shown in

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

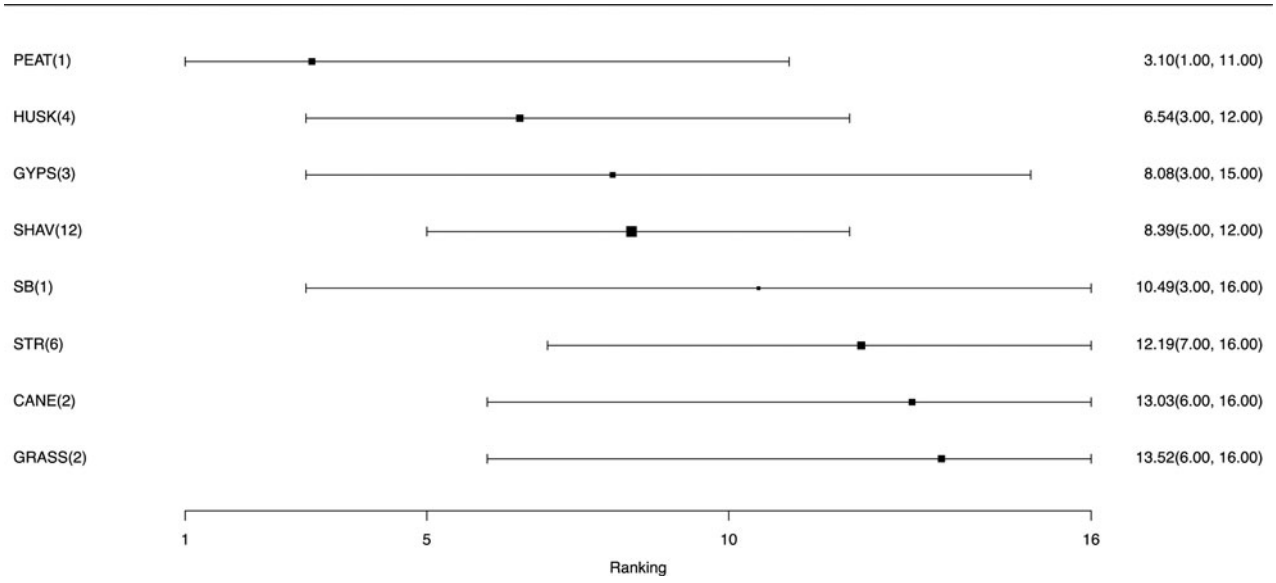

Fig. 6. Forest plot for ranking of intervention arms for litter management options evaluating the presence or absence of footpad lesions as an outcome for litter management in broiler chickens. Mean rank and 95% credibility intervals are shown.

Supplementary Tables S8 and S9, respectively. For each outcome, there were a considerable number of comparisons in which there were major concerns about the potential for bias under the domain of imprecision, reflecting the large confidence intervals resulting from the small number of trials contributing to each comparison. There were few concerns about bias due to heterogeneity; this result was expected based on the wide confidence intervals on the RRs, again due to the small number of studies contributing to any given comparison.

The contribution of studies to the RR based on the approach to randomization for each comparison between the litter management options was calculated, and the results are shown in Supplementary Figs. S4 and S5 for mortality and footpad lesions, respectively. For both mortality and presence or absence of footpad lesions as outcomes, a visual inspection of the resulting figures illustrates that a substantive component of the evidence for many intervention comparisons had a high (red) or unclear (yellow) risk of bias due to the randomization process.

The contribution of studies to the RR for each comparison based on blinding is not shown for either outcome; there were some concerns about the possible presence of bias in all comparisons, as blinding of outcome assessors was not reported in any of the trials.

Additional analyses: pairwise meta-analysis

The results of the pairwise meta-analysis comparing mortality risk for broilers raised on fresh litter versus used litter are shown in Fig. 7. The meta-analysis included 12 comparisons from seven trials (Aggarwal *et al.*, 1978; Jones and Hagler, 1982; Malone *et al.*, 1990; Malone and Gedamu, 1995; Balogun *et al.*, 1999; Vieira and Moran, 1999; Nunes *et al.*, 2012; Xu *et al.*, 2015; Shepherd *et al.*, 2017; Garcés-Gudiño *et al.*, 2018). The analysis indicated no association between fresh versus used litter on the risk of mortality (OR = 1.00, 95% CI = 0.84, 1.20), although there was considerable heterogeneity present ($I^2 = 66\%$). There was no compelling evidence of small study effects based on a visual assessment of the funnel plot (Fig. 8).

Discussion

Summary of evidence

A considerable body of work addressed the topic of litter management to reduce morbidity and mortality in broiler chickens. The NMAs conducted in the present study did not indicate differences in mortality risks or footpad lesions for poultry exposed to different litter types, flooring types, or litter additives type. In addition, the pairwise meta-analysis did not find differences in mortality risk between poultry housed on fresh versus used litter. However, these results should be interpreted cautiously because of the small number of trials investigating each of the specific interventions, with the exception of the shavings and husks as the type of litter. It is important that trials be replicated in order to ensure that the results of investigations into the effects of interventions on given outcomes are consistent and accurate. Replication is an essential need for any research synthesis, be it expert opinion, narrative review, or meta-analysis. The advantage of a quantitative approach, such as meta-analysis or NMA, is that results can provide a visual representation, such as that in the figure of the ranking plots, of the uncertainty in a body of work that is not communicated if the same body of evidence is summarized either by expert opinion or narrative review.

Limitations of the body of research

In the NMA, similar interventions were combined to create intervention categories; for instance, the intervention category 'paper' litter included different densities of paper, recycled versus new paper, as well as chips, particles, and pellets (data not shown). The decision to combine similar interventions was because these variations in factors such as paper depth were not considered to represent truly different interventions, and because there would have been little or no replication of interventions without some combination into categories. However, combining similar but non-identical interventions does introduce heterogeneity into the definition of the intervention, and therefore potentially into the results of the NMA. For both the mortality outcome

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

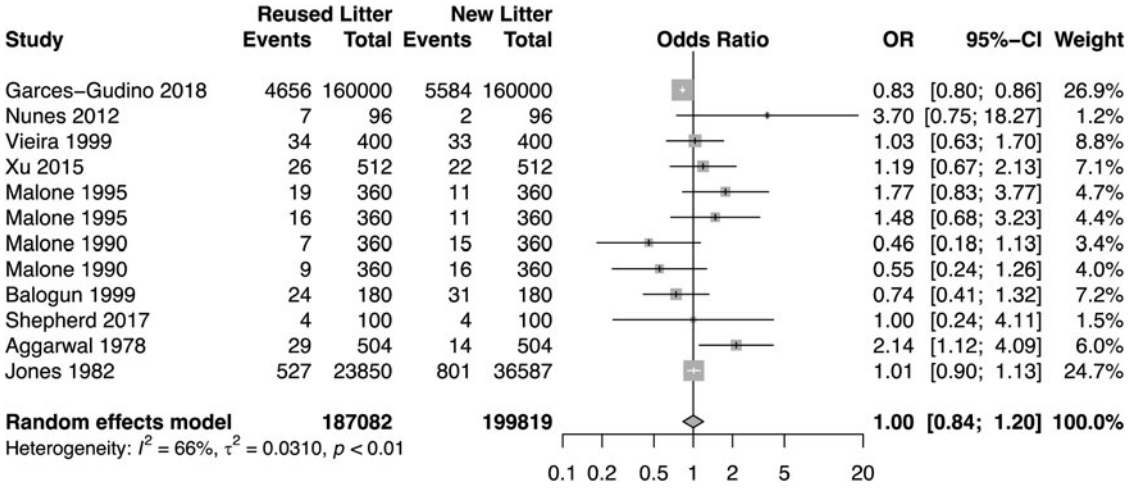

Fig. 7. Pairwise meta-analysis of fresh versus used litter to prevent mortality in broiler chickens.

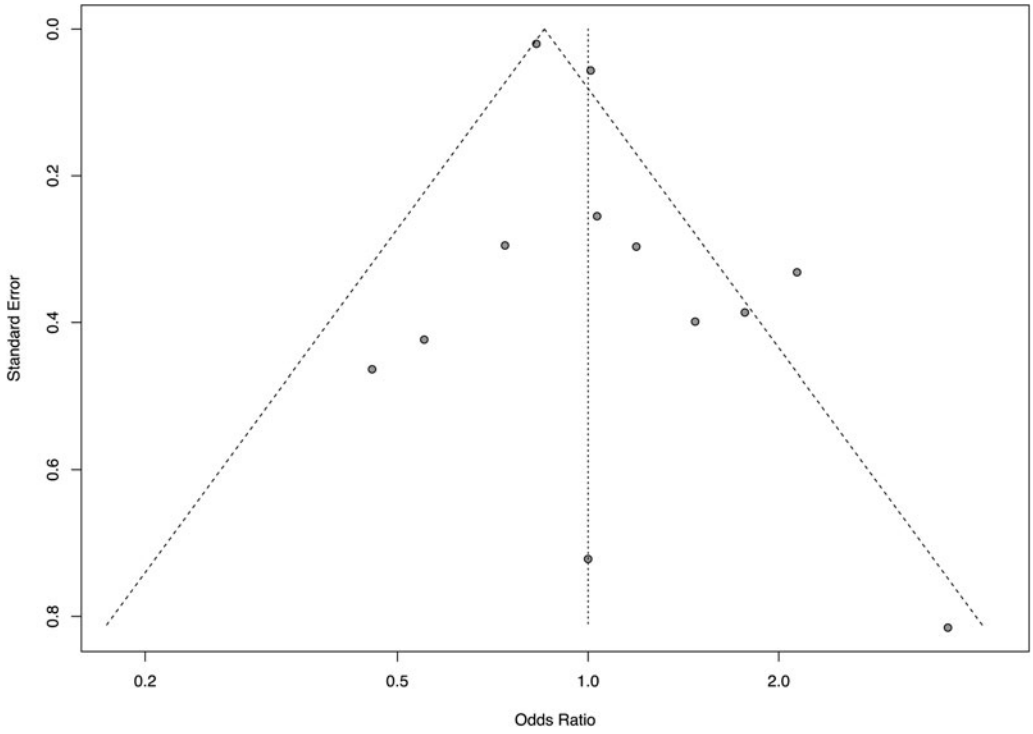

Fig. 8. Funnel plot for visual examination of small study effects from a pairwise meta-analysis of fresh versus used litter to prevent mortality in broiler chickens.

and the footpad outcome, there also was an intervention comparison that did not link into the network, meaning that these interventions were not common to any of the other intervention arms reported in the included trials. Researchers conducting future trials, and organizations funding research, could refer to the geometry and relative size of nodes in the networks developed in the present analysis to ensure that at least one of the interventions included in future trials are represented in the existing network. Additionally, researchers and research funders can use the information from the networks to target comparisons in future trials to address current gaps; this maximizes research efficiency by ensuring that future trials on litter management will build on the overall body of literature on this topic.

There was also considerable variation in the outcomes that were reported among trials, and in the ways in which the outcomes were measured (for example, using continuous versus binary measures). When synthesizing research results, continuous and binary outcomes cannot be combined, although it is possible to create a binary outcome from data presented on a scale if the value on the scale that corresponds to a zero score is identified. Variation in the outcomes that are reported across trials also reduces the power of evidence synthesis because different outcomes (e.g. breast blisters and footpad lesions) cannot be meaningfully combined in any research synthesis method. In the human healthcare literature, there are ongoing initiatives seeking to create core outcome sets for clinical trials that aim to address

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

the heterogeneity in the outcomes reported between trials; an example is the COMET (Core Outcome Measures in Effectiveness Trials) initiative (Williamson *et al.*, 2017). As examples, Wuytack *et al.* (2018) and O'Donnell *et al.* (2019) provide examples of recent protocols for developing core outcome sets. Creation of a core outcome set requires the input of content experts who conduct trials in the target area to determine which outcomes are of critical importance. However, the establishment of core outcome sets does not restrict researchers from reporting other outcomes in their trial reports; rather, the intention is to provide guidelines for a minimum set of outcomes to be included in all trials. Including at least some common outcomes in all trials can help to maximize the value, comparability, and synthesis potential of trials addressing the same issue. Creating core outcome sets for studies evaluating interventions to improve health and thereby reduce antibiotic use in poultry may be warranted.

None of the studies included in the NMAs incorporated adjustments for non-independence of birds (i.e. clustering) within flocks. All of the data from the included trials were presented as raw numbers (i.e. number of birds with and without the outcome of interest and a total number of birds with intervention group) or as univariable measures of association, yet all of the birds were housed in groups (i.e. flocks or barns). Therefore, the stated sample sizes in the included studies do not necessarily correspond to the effective sample sizes included in the analyses. Statistical methods to control for clustering are available in most statistical packages. When appropriate control for clustering is not incorporated in an analysis, the resultant confidence intervals will be inappropriately narrow (Schukken *et al.*, 2003). The lack of control for clustering should be considered when interpreting the results of this NMA, i.e. if clustering were controlled, the imprecision would be even greater.

In the literature identified for this review, there were issues related to the quality of reporting in the trials, including the reporting of study characteristics as well as key design features related to the risk of bias domains. Concerns related to the quality of reporting of trials in livestock populations have been identified in studies designed to evaluate reporting practices (Wellman and O'Connor, 2007; Burns and O'Connor, 2008; Sargeant *et al.*, 2009a, 2009b; Winder *et al.*, 2019). To address concerns about the quality of reporting, the REFLECT statement was developed by an expert consensus process to provide guidelines for the reporting of clinical trials in livestock populations (O'Connor *et al.*, 2010a). The RELFECT statement includes a 22-item checklist describing which components of a trial should be reported, as well as an explanation and elaboration document that provides further explanation on the importance of each item (Sargeant *et al.*, 2010a). Examples of trials reporting each item in a comprehensive manner are also provided. The REFLECT statement methods and elaboration documents were co-published in multiple journals (O'Connor *et al.*, 2010a, 2010b, 2010c, 2010d, 2010e; Sargeant *et al.*, 2010a, 2010b) and also are available online (http://www.reflect-statement.org/; https://meridian.cvm.iastate.edu/). Although not specific to poultry, the REFLECT statement provides relevant reporting guidelines for any trials conducted in animal populations housed in groups. Using the REFLECT statement to guide the writing of trial reports will increase the ability of the readers of these reports to assess the external validity and the generalizability of the trial results to their population of interest. Compliance with REFLECT guidelines will also help the readers to evaluate the internal validity or potential for bias in the trial. Ultimately, this will increase the utility of the research being conducted on litter management options, as well as other animal health research.

Limitations of the review

It is possible that not all relevant literature was captured by our search, particularly given the vertical nature of the poultry industry and the likelihood of internal research being conducted by poultry companies that might not be made available in the public realm. We also included only English-language articles, and the exclusion of 12 non-English articles during the full-text screening may have biased our results. Other relevant non-English-language articles may have been missed by the search, which incorporated only English search terms. We also collapsed intervention arms to increase the power of the analysis, and this may have impacted the results, if, for example, beneficial and harmful litter types were collapsed into a common intervention arm. Nonetheless, this approach was deemed necessary, as was little exact replication of interventions among studies.

Conclusions

We reviewed the literature related to the impacts of litter management options on poultry health; these management options included litter type, flooring type, litter additives, litter depth, and the use of fresh versus reused litter. Although there was a large number of studies, there was considerable variation in the specific interventions that were assessed, as well as in the outcomes that were measured in different trials. NMAs of the impacts of litter type, flooring type, or additives on mortality and on the presence or absence of footpad lesions did not reveal large differences in the outcomes among the interventions, and the pairwise meta-analysis did not indicate significant differences in mortality risks between poultry exposed to new and used litter. These results were based on a small number of comparisons among most interventions. In many cases, the trials identified in this review did not report the study characteristics and of the trial design features related to the potential for bias. Well-reported trials using at least one intervention that has previously been reported in the literature, as well as reporting outcomes that are common in the literature, would help to grow the knowledge base on the potential impact of litter management on bird health.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1466252319000227.

Author contributions. JMS developed the review protocol, coordinated the project team, assisted with the data analysis, interpreted the results, and wrote the manuscript drafts. MdB, KC, KD, BD, JD, SM, CM, MR conducted relevance screening, extracted data, conducted risk of bias assessments, commented on manuscript drafts, and approved the final manuscript. DH conducted the data analysis, provided guidance for the interpretation of the results, commented on manuscript drafts, and approved the final manuscript. AMOC, CML, and AV assisted with the development of the review protocol, provided guidance on the interpretation of the results, commented on manuscript drafts, and approved the final manuscript. YS and CW provided guidance on the interpretation of the analysis and results, commented on manuscript drafts, and approved the final manuscript. CBW assisted with the development of the review protocol, assisted with data screening, data extraction and risk of bias assessment, conducted the analysis, provided guidance on the interpretation of the results, commented on manuscript drafts, and approved the final manuscript draft.

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press

Financial support. Support for this project was provided by The Pew Charitable Trusts.

Conflict of interest. None of the authors has conflicts to declare.

References

Abreu VMN, de Abreu PG, Jaenisch FRF, Coldebella A and de Paiva DP (2011). Effect of floor type (dirt or concrete) on litter quality, house environmental conditions, and performance of broilers. *Brazilian Journal of Poultry Science* **13**, 127–137.

Aggarwal CK, Singh RA, Gupta SC and Pal RN (1978). Performance of broiler chickens raised on built up and fresh litters with their feed supplemented with antibiotics, coccidiostats and their combinations. *Indian Veterinary Journal* **55**, 952–957.

Akpobome GO and Fanguy RC (1992). Evaluation of cage floor systems for production of commercial broilers. *Poultry Science* **71**, 274–280.

Avdalovic V, Vucinic M, Resanovic R, Avdalovic J, Maslic-Strizak D and Vucicevic M (2017). Effect of pelleted and chopped wheat straw on the footpad dermatitis in broilers. *Pakistan Journal of Zoology* **49**, 639–1646.

Balogun TF, Adekeye JO, Olumeyan DB, Afolayan SB and Bawa GS (1999). Effect of change of litter material on voluntary water consumption, litter and faecal microbial load and performance of broiler chickens. *Bulletin of Animal Health and Production in Africa* **47**, 71–76.

Burns MJ and O'Connor AM (2008). Assessment of methodological quality and sources of variation in the magnitude of vaccine efficacy: a systematic review of studies from 1960 to 2005 reporting immunization with *Moraxella bovis* vaccines in young cattle. *Vaccine* **26**, 144–152.

Çavuşoğlu E, Petek M, Abdourhamane IM, Akkoc A and Topal E (2018). Effects of different floor housing systems on the welfare of fast-growing broilers with an extended fattening period. *Archives Animal Breeding* **61**, 9–16.

Chen Z and Jiang X (2014). Microbiological safety of chicken litter or chicken litter-based organic fertilizers: a review. *Agriculture* **4**, 1–29.

de Jong IC, Gunnink H and van Harn J (2014). Wet litter not only induces footpad dermatitis but also reduces overall welfare, technical performance, and carcass yield in broiler chickens. *Journal of Applied Poultry Research* **23**, 51–58.

Dias S, Welton NJ, Caldwell DM and Ades AE (2010). Checking consistency in mixed treatment comparison meta-analysis. *Statistics in Medicine* **29**, 555–560.

Dias S, Welton NJ, Sutton AJ and Ades AE (2011). NICE DSU Technical Support Document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. Technical Support Document in Evidence Synthesis, no. TSD2, National Institute for Health and Clinical Excellence. Available at https://research-information.bristol.ac.uk/en/publications/nice-dsu-technical-support-document-2-a-.generalised-linear-modelling-framework-for-pairwise-and-network-metaanalysis-of-randomised-controlled-trials(f941ebe0-73fa-444f-85d1-aa311d7b50f6)/e.

Do JC, Choi IH and Nahm KH (2005). Effects of chemically amended litter on broiler performances, atmospheric ammonia concentration, and phosphorus solubility in litter. *Poultry Science* **84**, 679–686.

Dunlop MW, Moss AF, Groves PJ, Wilkinson SJ, Stuetz RM and Selle PH (2016). The multidimensional causal factors of 'wet litter' in chicken-meat production. *Science of the Total Environment* **562**, 766–776.

European Food Safety Authority (2010). Application of systematic review methodology to food and feed safety assessments to support decision making. *European Food Safety Authority Journal* **8**, 1637–1727.

Garcés-Gudiño J, Merino-Guzmán R and Cevallos-Gordón AL (2018). Litter reuse reduces *Eimeria spp* oocyst counts and improves the performance in broiler chickens reared in a tropical zone in Ecuador. *European Poultry Science* **82**, 1–9.

Gholap D, Macklin K, Blake J and Bilgili S (2012). Estimation of the efficacy of a chemical and a biological litter amendment in reducing ammonia levels, bacterial load, and paw lesions in commercial broiler facilities. In *Western Poultry Disease Conference*, pp. 103–105.

Hafeez A, Suhail SM, Durrani FR, Jan D, Ahmad I, Chand N and Rehman A (2009). Effect of different types of locally available litter materials on the performance of broiler chicks. *Sarhad Journal of Agriculture* **4**, 581–586.

Higgins JPT and Green S (2011). *Cochrane Handbook for Systematic Reviews of Interventions*. West Sussex, England: John Wiley & Sons, Ltd.

Higgins JPT, Sterne JAC, Savović J, Page MJ, Hróbjartsson A, Boutron I, Reeves B and Eldridge S (2016). A revised tool for assessing risk of bias in randomized trials. *Cochrane Database of Systematic Reviews* **10**, 29–31.

Hu D, Wang C and O'Connor AM (2019). A method of back-calculating the log odds ratio and standard error of the log odds ratio from the reported group-level risk of disease. *bioRxiv*. doi.org/10.1101/760942.

Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP, Mulrow C, Catalá-López F, Gøtzsche PC, Dickersin K, Boutron I, Altman DG and Moher D (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Annals of Internal Medicine* **162**, 777–784.

Jones FT and Hagler WM (1982). Observations on new and reused litter for growing broilers. *Poultry Science* **62**, 175–179.

Kaukonen E, Norring M and Valros A (2016). Effect of litter quality on foot pad dermatitis, hock burns and breast blisters in broiler breeders during the production period. *Avian Pathology* **45**, 667–673.

Malone GW and Gedamu N (1995). Pelleted newspaper as a broiler litter material. *Journal of Applied Poultry Research* **4**, 49–54.

Malone GW, Tilmon HD and Taylor RW (1990). Evaluation of kenaf core for broiler litter. *Poultry Science* **69**, 2064–2067.

Mendes AS, Paixão SJ, Restelatto R, Reffatti R, Possenti JC, de Moura DJ, Morello GMZ and de Carvalho TMR (2011). Effects of initial body weight and litter material on broiler production. *Brazilian Journal of Poultry Science* **13**, 165–170.

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P and Stewart LA (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Review* **4**, 1.

Moura CAA, Totton SC, Sargeant JM, O'Sullivan TL, Linhares DCL and O'Connor AM (2019). Evidence of improved reporting of swine intervention trials in the post-REFLECT statement publication period. *Journal of Swine Health and Production* **27**, 265–277.

Nunes RV, Scherer C, Pozza PC, Eyng C, Bruno LDG and Vieites FM (2012). Use of probiotics to replace antibiotics for broilers. *Revista Brasileira de Zootecnia* **41**, 2219–2224.

O'Connor AM and Sargeant JM (2014). An introduction to systematic reviews in animal health, animal welfare and food safety. *Animal Health Research Reviews* **15**, 3–13.

O'Connor AM, Sargeant JM, Gardner I, Dickson J, Torrence M, Dewey CE, Dohoo I, Evans R, Gray J, Greiner M, Keefe G, Lefebvre S, Morley P, Ramirez A, Sischo W, Smith D, Snedeker, K, Sofos, J, Ward M and Wills R (2010a). The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety. *Zoonoses and Public Health* **57**, 95–104.

O'Connor AM, Sargeant JM, Gardner I, Dickson J, Torrence M, Dewey CE, Dohoo I, Evans R, Gray J, Greiner M, Keefe G, Lefebvre S, Morley P, Ramirez A, Sischo W, Smith D, Snedeker K, Sofos J, Ward M and Wills R (2010b). The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety. *Preventive Veterinary Medicine* **93**, 11–18.

O'Connor AM, Sargeant JM, Gardner I, Dickson J, Torrence M, Dewey CE, Dohoo I, Evans R, Gray J, Greiner M, Keefe G, Lefebvre S, Morley P, Ramirez A, Sischo W, Smith D, Snedeker, K, Sofos, J, Ward M and Wills R (2010c). The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety. *Journal of Veterinary Internal Medicine* **24**, 57–64.

O'Connor AM, Sargeant JM, Gardner I, Dickson J, Torrence M, Dewey CE, Dohoo I, Evans R, Gray J, Greiner M, Keefe G, Lefebvre S, Morley P, Ramirez A, Sischo W, Smith D, Snedeker K, Sofos J, Ward M and Wills R (2010d). The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety. *Journal of Food Protection* **73**, 132–139.

O'Connor AM, Sargeant JM, Gardner I, Dickson J, Torrence M, Dewey CE, Dohoo I, Evans R, Gray J, Greiner M, Keefe G, Lefebvre S, Morley P, Ramirez A, Sischo W, Smith D, Snedeker K, Sofos J, Ward M and Wills R (2010e). The REFLECT statement: methods and processes of creating reporting guidelines for randomized controlled trials for livestock and food safety. *Journal of Swine Health and Production* **18**, 18–26.

O'Connor AM, Coetzee JF, da Silva N and Wang C (2013). A mixed treatment comparison meta-analysis of antibiotic treatments for bovine respiratory disease. *Preventive Veterinary Medicine* **110**, 77–87.

O'Connor AM, Sargeant JM and Wang C (2014). Conducting systematic reviews of intervention questions III: synthesizing data from intervention studies using quantitative approaches (meta-analysis). *Zoonoses and Public Health* **61**, 52–63.

O'Donnell CM, Black N, McCourt KC, McBrien ME, Clarke M, Patterson CC, Blackwood B, McAuley DF and Shields MO (2019). Development of a core outcome set for studies evaluating the effects of anaesthesia on perioperative morbidity and mortality following hip fracture surgery. *British Journal of Anaesthesia* **122**, 120–130.

Papakonstantinou T, Nikolakopoulou A, Rücker G, Chaimani A, Schwarzer G, Egger M and Salanti G (2018). Estimating the contribution of studies in network meta-analysis: paths, flows and streams. *F1000Research* **7**, 610.

Petek M, Çibik R, Yildiz H, Sonat FA, Gezen SS, Orman A and Aydin C (2010). The influence of different lighting programs, stocking densities and litter amounts on the welfare and productivity traits of a commercial broiler line. *Veterinarija Ir Zootechnika* **51**, 36–43.

Plummer, M (2015). RJAGS: Bayesian graphical models using MCMC. R. Package version 3.15. Available at http://CRAN.R-project.org/package=rjags (Accessed 15 April 2019).

R Core Team (2015). R: A language and environment for statistical computing. R Foundation For Statistical Computing. Available at https://www.R-project.org (Accessed 15 April 2019).

Sahoo SP, Kaur D, Sethi APS, Sharma A, Chandra M and Chandrahas AS (2017). Effect of chemically amended litter on litter quality and broiler performance in winter. *Journal of Applied Animal Research* **45**, 533–537.

Salanti G (2012). Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods* **3**, 80–97.

Salanti G, Del Giovane C, Chaimani A, Caldwell DM and Higgins JPT (2014). Evaluating the quality of evidence from a network meta-analysis. *PLoS ONE* **9**, e99682.

Sargeant JM and O'Connor AM (2014). Introduction to systematic reviews in animal agriculture and veterinary medicine. *Zoonoses and Public Health* **61**, 3–9.

Sargeant JM, Elgie R, Valcour J, Saint-Onge J, Thompson A, Marcynuk P and Snedeker K (2009a). Methodological quality and completeness of reporting in clinical trials conducted in livestock species. *Preventive Veterinary Medicine* **91**, 107–115.

Sargeant JM, Saint-Onge J, Valcour J, Thompson A, Elgie R, Snedeker and Marcynuk JP (2009b). Quality of reporting in clinical trials of pre-harvest food safety interventions and associations with treatment effect. *Foodborne Pathogens and Disease* **6**, 989–999.

Sargeant JM, O'Connor AM, Gardner IA, Dickson JS, Torrence ME and Consensus Meeting Participants: Dohoo IR, Lefebvre SL, Morley PS, Ramirez A and Snedeker K (2010a). The REFLECT statement: reporting guidelines for randomized controlled trials in livestock and food safety: explanation and elaboration. Journal of Food Protection 73, 579–603.

Sargeant JM, O'Connor AM, Gardner IA, Dickson JS, Torrence ME and Consensus Meeting Participants: Dohoo IR, Lefebvre SL, Morley PS, Ramirez A and Snedeker K (2010b). The REFLECT statement: reporting guidelines for randomized controlled trials in livestock and food safety: explanation and elaboration. Zoonoses and Public Health 57, 105–136.

Schukken YH, Grohn YT, McDermott B and McDermott JJ (2003). Analysis of correlated discrete observations: background, examples and solutions. *Preventive Veterinary Medicine* **59**, 223–240.

Schwarzer G (2019) 'Meta.' Journal Article. CRAN, Version 4.94.

Shepherd EM and Fairchild BD (2010). Footpad dermatitis in poultry. *Poultry Science* **89**, 2043–2051.

Shepherd EM, Fairchild BD and Ritz CW (2017). Alternative bedding materials and litter depth impact litter moisture and footpad dermatitis. *Journal of Applied Poultry Research* **26**, 518–528.

Sterne JAC, Gavaghan D and Egger M (2000). Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology* **53**, 1119–1129.

Vargas-Galicia AJ, Sosa-Montes E, Rodríguez-Ortega LT, Pro-Martinez A, Ruiz-Feria CA, González-Cerón F, Gallegos-Sánchez J, Arreola-Enríquez J and Bautista-Ortega J (2017). Effect of litter material and stocking density on bone and tendon strength, and productive performance in broilers. *Canadian Journal of Animal Science* **97**, 673–682.

Viechtbauer W (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**, 1–48.

Vieira SL and Moran ET (1999). Effects of delayed placement and used litter on broiler yields. *Journal of Applied Poultry Research* **8**, 75–81.

Wellman NG and O'Connor AM (2007). Meta-analysis of treatment of cattle with bovine respiratory disease with tulathromycin. *Journal of Veterinary Pharmacology and Therapeutics* **30**, 234–241.

White IR, Barrett JK, Jackson D and Higgins JPT (2012). Consistency and inconsistency in network meta-analysis: model estimation using multiple meta-regression. *Research Synthesis Methods* **3**, 111–125.

Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, Clarke M, Gargon E, Gorst S, Harman N and Kirkham JJ (2017). The COMET handbook: version 1.0. *Trials* 18, 280.

Winder CB, Churchill KJ, Sargeant JM, LeBlanc SJ, O'Connor AM and Renaud DL (2019). REFLECTing on a year of animal trials in the journal of dairy science. *Journal of Dairy Science* **102**, 4759–4771.

World Health Organization (2015). Global action plan on antimicrobial resistance. Available at http://apps.who.int/iris/bitstream/handle/10665/193736/9789241509763_eng.pdf?sequence=1. (Accessed 3 January 2019).

Wuytack F, Gutke A, Stuge B, Mørkved S, Olsson C, Robinson HS, Vøllestad NK, Öberg B, Wikmar LN, Mena JJS and Smith V (2018). Protocol for the development of a core outcome set for pelvic girdle pain, including methods for measuring the outcomes: the PGP-COS study. *BMC Medical Research Methodology* **18**, 158.

Xu Y, Stark CR, Ferket PR, Williams CM, Nusairat B and Brake J (2015). Evaluation of litter type and dietary coarse ground corn inclusion on broiler live performance, gastrointestinal tract development, and litter characteristics. *Poultry Science* **94**, 362–370.

Yegani M and Korver DR (2008). Factors affecting intestinal health in poultry. *Poultry Science* **87**, 2052–2063.

https://doi.org/10.1017/S1466252319000227 Published online by Cambridge University Press