The Old School, Brewhouse Hill, Wheathampstead,
Hertfordshire AL4 8AN, UK
www.ufaw.org.uk

Animal Welfare 2018, 27: 263-274
ISSN 0962-7286
doi: 10.7120/09627286.27.3.263

Development and refinement of three animal-based broiler chicken welfare indicators

APO Souza†, VS Soriano†, MA Schnaider†, DS Rucinque‡ and CFM Molento*†

† Animal Welfare Laboratory, Federal University of Paraná, Rua dos Funcionários 1540, 80035-050, Curitiba, Paraná, Brazil
‡ Faculty of Animal Science and Food Engineering, University of São Paulo, Av Duque de Caxias Norte 225, 13635-900, Pirassununga, São Paulo, Brazil
* Contact for correspondence and requests for reprints: carlamolento@ufpr.br

Abstract

*This study aimed to refine bird-soiling as a broiler chicken (*Gallus gallus domesticus*) welfare indicator, and to develop and test two additional indicators, namely contact dermatitis on the breast and abdominal areas and carcase scratches. We constructed a questionnaire with pictures of birds presenting different indicator levels for classification as absent, low, moderate or severe. The questionnaire was sent to 146 invited experts for the first round and 88 for the second, in a Delphi process. Visual scales were built for the target indicators, which were tested by three assessors in ten flocks on-farm (n = 1,303 birds) and at the slaughterhouse (n = 1,631 birds). High concordance was observed among groups of Delphi respondents and among assessors. A total of 90.7% of the birds were either moderately or severely soiled, 99.9% were poorly feathered, 73.4 and 90.0% presented erythaema and carcase scratches, respectively. The correlations between litter quality and all outcomes assessed on-farm, and between bird-soiling and contact dermatitis on the breast and abdominal areas, were moderate. Results suggest that adoption of the proposed scales may improve our ability to assess broiler chicken welfare, since relevant problems were prevalent and measurement consistency acceptable. Substantial concordance observed among assessors encourages application of these animal-based indicators to assess broiler chicken welfare in a wide range of poultry houses, in a variety of different countries, thereby allowing the scales to be tested in a host of animal welfare conditions.*

Keywords: *animal welfare, outcomes, plumage cleanliness, skin irritation, welfare assessment, welfare measures*

Introduction

Animal welfare assessment may include animal- and resource-based indicators. The use of animal-based indicators to assess animal welfare has been encouraged (Veissier *et al* 2008; Rushen *et al* 2011; European Food Safety Authority [EFSA] 2012; World Organisation for Animal Health [OIE] 2013) and applied for regulatory purposes (European Commission 2017). The Welfare Quality® project proposed to standardise animal welfare assessment through the application of predominantly animal-based, scientifically validated measures (Blokhuis *et al* 2010) and it has been considered a robust tool to assess animal welfare (Webster 2009). The protocol for poultry includes measures of welfare related to four principles, ie good feeding, good housing, good health and appropriate behaviour (Welfare Quality® 2009).

Recent studies applying the Welfare Quality protocol® (2009) to assess broiler chicken (*Gallus gallus domesticus*) welfare suggested a need for refinement of some animal welfare measures. For example, Federici *et al* (2016) reported difficulties in assessing plumage cleanliness using the Welfare Quality protocol® scoring system, since birds assessed on-farm were poorly feathered on the breast. The visual eight-point scale developed by Wilkins *et al* (2003) to assess plumage cleanliness at the slaughterhouse was transformed in a four-point scale in the Welfare Quality protocol® to be assessed on-farm. However, type of soiling and bird feathering presented in the protocol pictures are not representative of the conditions observed in commercial farms for fast-growing broiler chickens. Previous studies have assessed bird cleanliness (Weeks *et al* 1994; Elwinger 1995; Dawkins *et al* 2004); however, details on the method used were often missing (Arnould *et al* 2009), suggesting the need for an updated scoring system.

A possible shortcoming in current broiler welfare assessment protocols is the absence of an effective measurement for contact dermatitis in the ventral body area for broiler chicken flocks, especially considering the high prevalence of contact dermatitis reported for other body parts, such as the foot-pads and hocks (Souza *et al* 2015; Tuyttens *et al* 2015; Federici *et al* 2016). The Welfare Quality protocol® (2009) provides a scoring system to assess the presence or absence of breast blisters. According to Greene *et al* (1985), flocks showing a high prevalence of foot-pad dermatitis are expected to also present other forms of contact dermatitis.

Contact dermatitis is an inflammatory skin reaction caused by contact with an offending substance (Muller 2001). Patches of erythaema and papules represent primary lesions and, as the inflammatory changes progress, crusting develops (Muller 2001). However, breast lesion scoring in broiler chickens has, instead, tended to focus on hyperkeratosis and ulcerations (Stephenson *et al* 1960; Greene *et al* 1985; Allain *et al* 2009; Gouveia *et al* 2009; Saraiva *et al* 2016). Contact dermatitis leads to poor animal welfare because of the pain caused by the lesions and the inflammatory process (Berg 2004). De Jong *et al* (2014) published the first study including the observation of erythaema on the breast during broiler chicken welfare assessment, evidencing the occurrence of primary signs of contact dermatitis on commercial farms. Thus, there is a need to develop a uniform scoring system, including the earlier stages of development of contact dermatitis on the breast.

Some broiler chicken production procedures, for example, moving birds from one place to another and exposing them to non-familiar human beings may increase fear and distress (Jones & Roper 1997; Jones *et al* 2002). In these cases, the assessment of carcase scratches seems to be a relevant welfare indicator (de Jong *et al* 2014), since birds experiencing panic and escape attempts frequently pile on top of each other, causing body lesions (Waiblinger *et al* 2006). Different scoring systems to assess carcase scratches have been used (Hargis *et al* 1989; Elfadil *et al* 1996; Pilecco *et al* 2012; Allain *et al* 2013), but no visual scale has been presented in previous studies on broilers. Considering that carcase scratches are not included in current broiler welfare assessment protocols, and that the lack of a standardised scoring method prevents comparison between studies, it seems important to further explore the potential use of this indicator.

The three indicators previously reviewed could be valuable contributors to existing welfare assessment schemes. This study aimed to refine bird-soiling (BS) as a broiler chicken welfare indicator, and to develop and test two additional indicators, namely contact dermatitis on the breast and abdominal areas (CD) and carcase scratches (CS).

Materials and methods

The first part of our study consisted of employing the Delphi method to develop the visual scales for the three indicators, with a basic description of each level of severity. Scales were then tested on-farm and at the slaughterhouse in order to study inter-rater reliability as well as correlation with other animal welfare indicators (AWI).

Delphi methodology

Three broiler chicken farms were visited in January 2016 to take pictures of birds presenting different levels of BS and CD. Birds were male and female Cobb 500® broilers, at 37 and 42 days of age. Carcases with different levels of scratches were sampled in a slaughterhouse, immediately prior to chilling. Pictures were analysed by one experienced researcher who pre-classified them as examples of absent, low, moderate and severe levels of each indicator. At least two pictures representative of each level to be included in the questionnaire were selected. The online questionnaire (https://www.onlinepesquisa.com) was developed in Portuguese and English and tested by three senior academic researchers in Brazil, the UK and Spain, with a deep knowledge of broiler chicken welfare and/or production.

The Delphi technique, ie a process of obtaining consensus from a group of expert respondents (Dajani *et al* 1979; Hsu & Sandford 2007), was closely adhered to. The questionnaire was sent to 146 experts in March 2016. Respondents were selected based on their experience and/or publications related to animal welfare and/or broiler chicken production, including university lecturers and researchers, professionals from the government and from private initiative represented by the meat industry and animal welfare certification schemes.

The questionnaire was divided by indicator, with the first ten pictures showing different examples of BS, followed by nine pictures of CS and ten of CD. On the first page of each indicator, all pictures were presented to familiarise the respondent with the range of picture variation. Thereafter, each picture was presented individually, and respondents asked to choose the best descriptor from absent, light, moderate and severe. Respondents could also give another descriptor in an open-ended text box. For each picture, a short explanatory text on the indicator was provided (see the supplementary material to papers published in *Animal Welfare* section on the UFAW website; https://www.ufaw.org.uk/the-ufaw-journal/supplementary-material). As respondents may not have been familiar with all the indicators, we added the alternative 'prefer not to answer this question'. A field to justify the score was provided. One picture of a soiled bird was repeated and respondents were asked whether poor feathering affects BS score. If the answer was yes, respondents were presented with four options that included: (i) to propose a mathematical model for BS that considers general feathering; (ii) to propose a model that considers the proportion of body area presenting poor feathering; (iii) to consider the worst BS score when poor feathering is observed; and (iv) other.

In the first round, respondents not answering all questions of at least one indicator, or those whose answers were considered inconsistent were excluded. Data were analysed using descriptive statistics. Following the method proposed by Rayens and Hahn (2000), the interquartile deviation (IQD) for each picture was calculated to verify consensus among respondents. When IQD = 0 or IQD = 1 with a proportion equal to or higher than 60% in one level, consensus was deemed adequate for that picture; when IQD = 1 with a proportion lower than 60% in one level or IQD > 1, consensus was not adequate, and the picture was considered for inclusion in the second round. Groups of respondents were tested using the Kendall's coefficient of concordance corrected for ties (*Wt*).

In the second round, a new questionnaire was sent to the 88 respondents who had completed the first round, in September 2016. It included a preview of the main results from the first round and new questions to further study each

https://doi.org/10.7120/09627286.27.3.263 Published online by Cambridge University Press

Table 1 Indicators and definitions for the assessment of broiler chicken welfare.

Indicator	Place	Sampling
Contact dermatitis on the breast and abdominal areas (CD)	On-farm	Visual inspection of 130 birds in five locations in each poultry house, following the scale developed in this study
Bird-soiling (BS)	On-farm	Visual inspection of 130 birds in five locations in each poultry house, following the scale developed in this study
Foot-pad dermatitis (FPD)	On-farm	Visual inspection of 130 birds in five locations in each poultry house, following a 5-point scale†
Hock burn (HB)	On-farm	Visual inspection of 130 birds in five locations in each poultry house, following a 5-point scale†
Feathering condition (FC)	On-farm	Visual inspection of ventral body area of 130 birds in five locations in each poultry house, according to the following 3-point scale: good feathering (complete or nearly complete feathering), moderate feathering (one or more featherless area < 5 cm in diameter), poor feathering (at least one featherless area ≥ 5 cm in diameter)‡
Litter quality	On-farm	Visual inspection of six locations in each poultry house, following a 5-point scale†
Carcase scratches (CS)	Slaughterhouse	Visual inspection on slaughter line of 130 birds per flock after plucking, following the scale developed in this study

† According to the Welfare Quality® (2009) for broiler meat chickens;
‡ Adapted from Welfare Quality® (2009) for laying hens.

indicator. For BS, again, a question related to the correlation between poor feathering and soiling was included. In CD, we presented two scales, including erythaema, based on justification given by respondents for each level of severity in the first round. The three-point scale included absence of erythaema, intermediate (levels light and moderate together, since there was no consensus within these levels in the first round; details in supplementary material; https://www.ufaw.org.uk/the-ufaw-journal/supplementary-material) and severe erythaema. The four-point scale included erythaema and the presence of brown spots and breast blisters, adapted from De Jong *et al* (2014). Our intention was to observe whether erythaema should be assessed separately or be part of a scale including other breast lesions caused by prolonged contact with litter. In CS, we asked participants to quantify their maximum accepted levels according to depth and length of lesion for each category, and we considered median values reported for each point. We also asked participants whether old scratches should be considered during the assessment. After the second round, a visual scale was defined for each indicator, including a basic description of each level of severity.

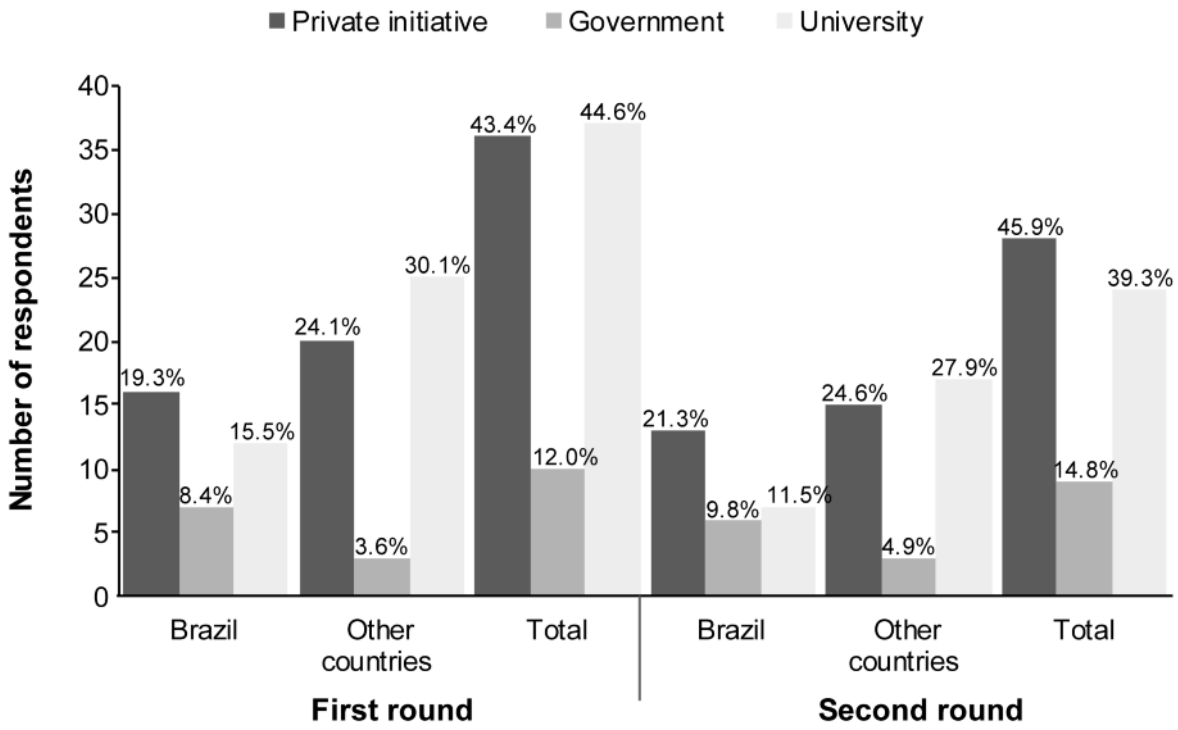
Testing of the visual scales on-farm and at the slaughterhouse

The testing of visual scales was performed on-farm in January 2017 (25°17'49.1"S, 54°05'41.7"W) and at the slaughterhouse in April of the same year (24º 55' 04" S, 50º 05' 50" W) in the State of Paraná, Southern Brazil, for each case in ten flocks. The sampling size of 1,300 birds was calculated considering a maximum error of 5 and 95% confidence interval. Two veterinarians (APOS and MAS) and one animal scientist (VSS) performed all assessments on-farm and at the slaughterhouse; APOS is experienced in auditing poultry farms and slaughterhouses. Assessors scored the animals simultaneously but independently. They underwent training, initially via picture observation, to learn how to assess each indicator, as shown in Table 1. The second step involved a training session at the Federal University of Paraná farm and in a slaughterhouse.

Two barns were visited each day, the first from 0800 to 1200h and the second from 1300 to 1700h. Poultry barns had sidewalls with wire mesh covered by blackout curtains working as a dark house (n = 1) and covered by yellow curtains, with natural lighting (n = 9), chosen as convenient samples according to our objective, to observe ranges of variation for each indicator rather than describing or comparing specific barn types or other factors. Birds were male and female Cobb 500®, assessed at 41.3 (± 2.0) days of age and weighing 2,147.3 (± 99.5) g at 35 days. All units had automatic feeders, nipple drinkers, sprinklers, exhaust fans and wood-shaving litter; nine units maintained evaporative cooling systems. Indoor mean temperature in the units at time of the visit was 27.7 (± 1.4)°C. Average broiler house area was 1,540 (± 187) m^2 and the number of birds per house was 18,904 (± 2,604), with a stocking density of 36.4 (± 0.9) kg m^{-2}. At the slaughterhouse, birds were Cobb 500®, Hubbard H1 and Ross 408, assessed at 27 (± 1) days of age and weighing 1,354 (± 35) g. To collect data on-farm a questionnaire was developed on the QuickTapSurvey® website and made available on a cell phone application to be used offline. Data from QuickTapSurvey® were downloaded into an Excel® database and checked for errors prior to use.

During training at the slaughterhouse to assess CS, initially, the sampling procedure described in the Welfare Quality® protocol (2009) for injuries, bruising and wing damage was followed, which demanded the observation of carcases passing the line for five to ten minutes. As the occurrence of scratches was higher than the occurrence of injuries, observation of all carcases was not feasible for the paired sampling required in our study. Thus, a specific

Figure 1



Number, origin and category of respondents in first (n = 83) and second (n = 61) rounds of the Delphi questionnaire on three broiler chicken welfare indicators, from March to October 2016.

procedure was performed to allow for the line speed: one assessor randomly established one carcase for every 8, 9 or 10 on the slaughter line to be evaluated by assessors simultaneously. This skipping method allowed for the assessment to be performed at a slower rhythm as compared to line speed and it was needed for adequate assessment and synchronisation across assessors.

Statistical analysis

Reliability between assessors was tested using Kendall's coefficient of concordance corrected for ties. Based on our perception during assessment, concordance was also tested by grouping light and moderate levels of CD; and moderate and severe levels of BS. Bird-soiling did not present ordinal distribution after grouping; thus, concordance was tested using the Fleiss' kappa coefficient. Coefficients from 0.61 to 0.80 were considered as substantial concordance, and from 0.81 to 1.0, almost perfect concordance (Landis & Koch 1977). The proportion of identical answers was calculated for feathering condition. Data were tested for normality using the Henze-Zirkler test followed by Spearman's rank correlation coefficient to test correlation between indicators. Correlation from 0.3 to 0.6 was considered moderate, and higher than 0.6 considered high (de Jong *et al* 2015). Analyses were performed using R Statistical Computing Environment software version 3.3.1 (R Core Team 2016).

Ethical approval

This project was approved by the Human Research Ethics Committee of the Health Science Sector (n 1,377,497; December 21st, 2015) and by the Animal Use Ethics Committee of the Agricultural Campus (n 079/2015; November 12th, 2015), both of the Federal University of Paraná.

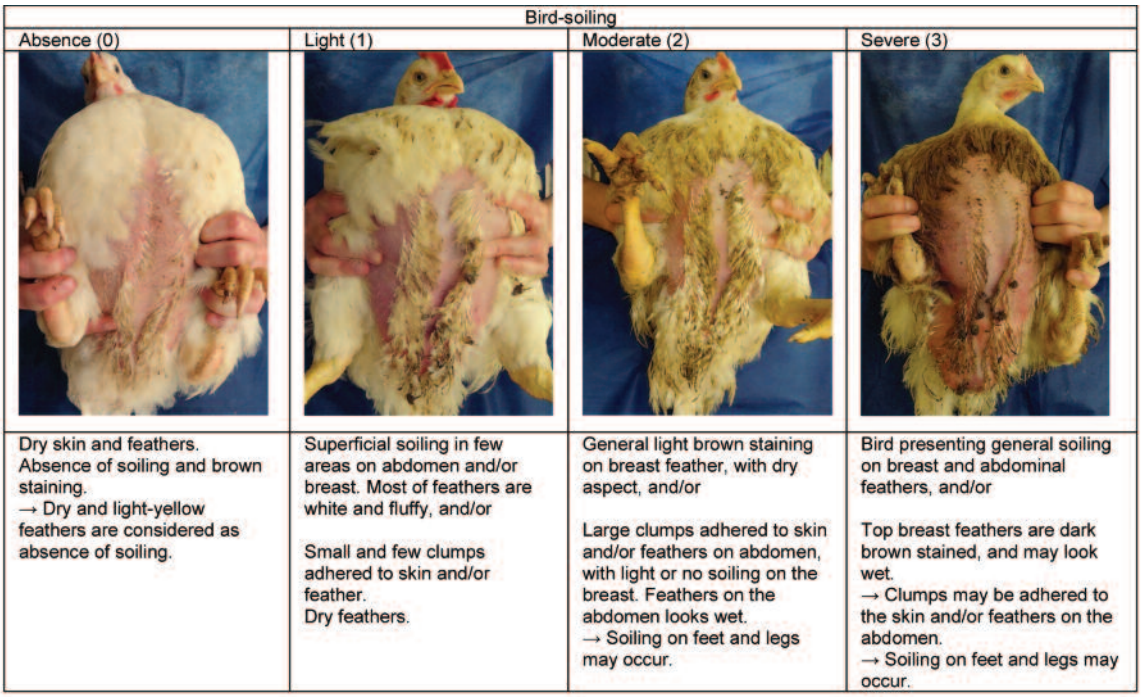
Results

Delphi methodology

In the first round, 60.3% (88/146) of invited experts answered the questionnaire. There were 56.8% (83/146) complete and relevant responses for BS, 55.5% (81/146) for CD, and 56.1% (82/146) for CS. In the second round, a total of 73.5% (61/88) of experts participated, and 68.7% (57/88) completed the questionnaire. Origin and number of respondents, presented as first followed by second round, were Brazil (35, 26), USA (14, 11), Canada (13, 10), UK (7, 6), Germany (3, 3), Belgium (2, 2), Sweden (2, 2), The Netherlands (2, 1), France (2,0), Italy (2,0) and Chile (1, 0). Proportion of respondents by category is presented in Figure 1. Answers were highly correlated among groups of respondents ($P < 0.001$; $Wt = 0.916$).

https://doi.org/10.7120/09627286.27.3.263 Published online by Cambridge University Press

Figure 2



Visual and descriptive scale to assess bird-soiling on-farm, developed using the Delphi methodology, from March to October 2016.

Bird-soiling

Consensus was achieved for 8/10 pictures in the first round. The visual scale presented in Figure 2 was constructed to be applied to broiler chicken farms. The question about the relation between feathering condition and BS score did not reach consensus in the first round. Respondents indicated that poor feathering affected BS assessment (57.8%; 48/83), most of them justifying that dirt appears to adhere more on feathers than on skin, thus, lack of feathers prevents clumps from being retained. A total of 36.1% (30/83) considered there to be no relation between BS and feathering, and 6.1% (5/83) gave other answers. In the second round, one justification representative of each main theme cited by respondents in the first round was presented, both for and against the relationship between feathering condition and BS, and respondents were asked to think again about this relationship. Results differed from the first round, and 83.6% (51/61) of respondents considered that poor feathering affected BS assessment. There was no consensus about the best option to integrate BS and poor feathering scores (see supplementary material; https://www.ufaw.org.uk/the-ufaw-journal/supplementary-material). Results clearly indicated that feathering is an issue and should be considered when assessing broiler chicken welfare. Based on this, feathering condition assessment during the on-farm testing of visual scale was included to further study this indicator (see Table 1).

Contact dermatitis on the breast and abdominal areas

Consensus was achieved for 2/10 pictures, being the most extreme cases, absence and severe CD. In the first round, 3.5% (3/83) of respondents did not consider the pictures representative of CD, and in the second round, a total of 64.9% (37/57) of respondents chose the four-point scale, most of them (62.2%, 23/37) because of the highest level of detail and information provided by the scale. A total of 31.6% (18/57) of respondents chose the three-point scale, most of them (61.1%, 11/18) justifying it was more practical and simpler. The scale presented in Figure 3 was chosen to be applied on-farm.

Carcase scratches

During the first round, consensus was achieved for 5/8 pictures. Although information was collected for descriptions of the four proposed levels of CS, answers were too generalised, using terms such as 'multiple scratches in one side', 'large area affected' and 'there are some deep scratches'. Answers were based on the following items: area affected (uni- or bilateral), quantity, depth (light or deep), length (small or long) and age (new or old) of scratches. In the second round, we aimed to quantify these items to establish clear thresholds between each CS level.

In the first round, respondents spontaneously presented different justifications based on age of scratches: some respondents considered recent scratches as more severe,

Figure 3

Contact dermatitis on the breast and abdominal areas			
Absence (0)	Light (1)	Moderate (2)	Severe (3)
			
Skin without lesion, inflammation or erythaema.	Light pink local- or generalised erythaema, or Skin may present different degrees of erythaema, from dark pink to red (up to 25% of ventral body area).	Skin may present different degrees of erythaema, from dark pink to red (from 25% to 50% of ventral body area), or Presence of small brown spots.	Large area inflamed (more than 50% of ventral body area), dark pink or red colour, or Large brown spots or breast blister.
		If moderate or severe, are there brown spots or breast blisters? ( )Yes ( )No	

Visual and descriptive scale to assess contact dermatitis on the breast and abdominal areas on-farm, developed using the Delphi methodology, from March to October 2016.

Figure 4

Carcase scratch (back, lateral and thigh)			
Absence (0)	Light (1)	Moderate (2)	Severe (3)
			
Intact skin, no visible scratches.	Up to 3 small or long superficial scratches combined, and/or Up to 1 small or long deep scratch. → Scratches may be uni- or bilateral; new or old.	Up to 5 small or long superficial scratches combined, and/or Up to 2 small or long deep scratches. → Scratches may be uni- or bilateral; new or old.	More than 5 small and long superficial scratches combined, and/or 3 or more small or long deep scratches. → Scratches may be uni- or bilateral; new or old.
- Small scratches: up to 3 cm (1.2 inches); long scratches: larger than 3 cm (1.2 inches) - Superficial scratches: light scratches, affecting the epidermis; deep scratches: affect the dermis, they may achieve muscle tissue - Old scratches: brown or yellow scratches, indicative of healing process; new scratches: pink or red scratches			

Visual and descriptive scale to assess carcase scratches at the slaughterhouse, developed using the Delphi methodology, from March to October 2016.

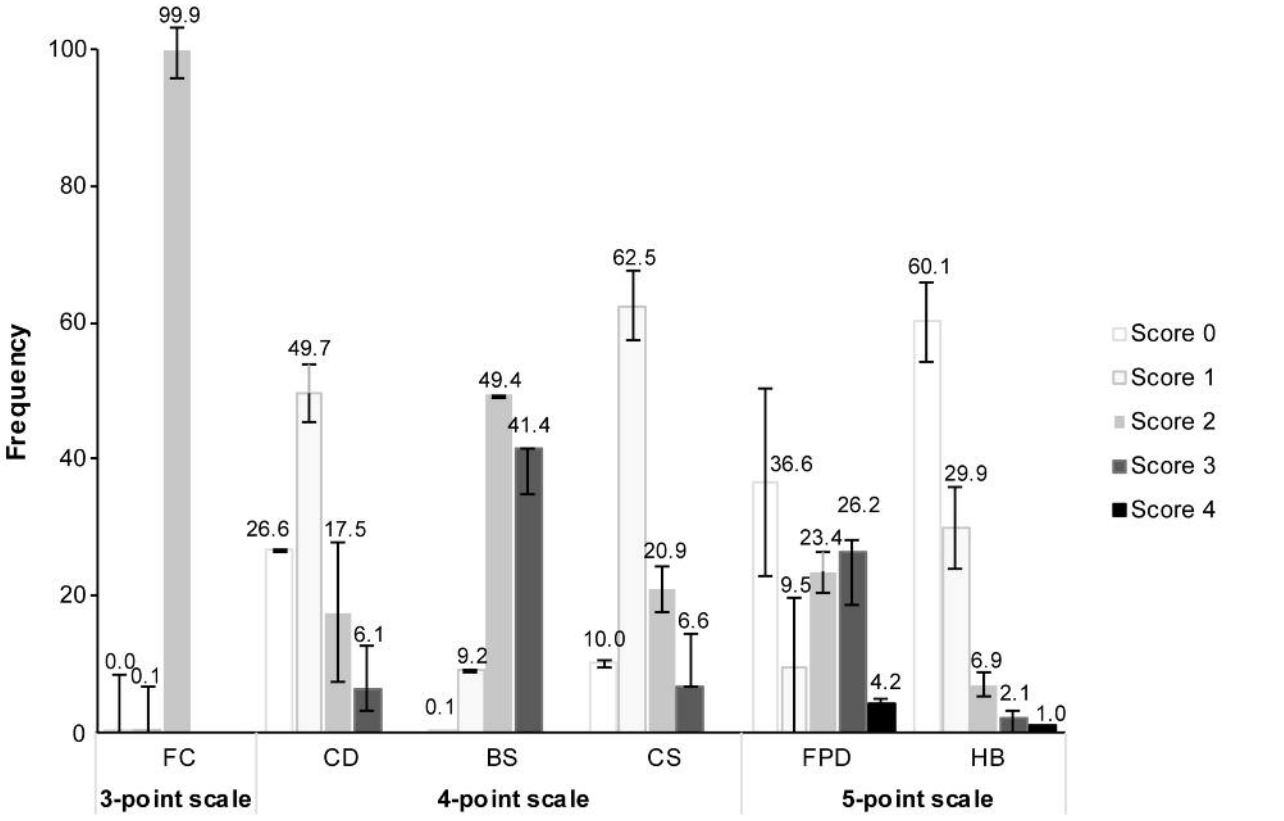
https://doi.org/10.7120/09627286.27.3.263 Published online by Cambridge University Press

Table 2 Level of concordance among three assessors, and correlation of broiler chicken welfare indicators measured on-farm or at the slaughterhouse, January and April 2017.

Indicator	Concordance among assessors		Correlation between indicators			
	Kendall's coefficient of concordance	Fleiss' kappa	Spearman rank correlation			
	A1×A2×A3	A1×A2×A3	BS	FPD	HB	Litter
Bird-soiling (BS)	0.739*			0.08*	0.25*	0.43*
BS scores 2+3	–	0.334**				
Contact dermatitis on the breast and abdominal areas (CD)	0.781*		0.34*	0.06*	0.24*	0.33*
CD scores 1+2	0.709*					
Foot-pad dermatitis (FPD)	0.941**				0.17*	0.35*
Hock burn (HB)	0.76**					0.31*
Carcase scratches (CS)†	0.74**					

A1, assessor 1; A2, assessor 2; A3, assessor 3.
* $P < 0.0001$; ** $P < 0.00001$.
† at the slaughterhouse.

Figure 5



Mean frequency of six broiler chicken welfare indicators measured by three assessors in 1,303 birds on-farm (ten flocks) and in 1,631 birds at the slaughterhouse (ten flocks). Poor feathering (FC) ranging from 0 (absence) to 2 (severe); contact dermatitis on the breast and abdominal areas (CD), bird-soiling (BS) and carcase scratches (CS) ranging from 0 (absence) to 3 (severe); foot-pad dermatitis (FPD) and hock burn (HB) ranging from 0 (absence) to 4 (severe).

some did not consider old scratches as a welfare problem, and some were concerned about the presence of old and new scratches simultaneously. As it was a new subject, in the second round, answers given by respondents about age of scratches were presented and opinions sought. As a result, 98.2% (56/57) of respondents considered old scratches should be assessed as an AWI. Most of them (89.3%, 50/56) clearly stated CS was a welfare problem regardless of when it occurred. Other respondents also included a justification based on economic loss due to slaughter condemnation of scratched carcases (12.5%, 7/56) and food safety concerns (3.6%, 2/56). Considering the first and the second rounds, the scale presented in Figure 4 was developed to be applied at the slaughterhouse.

Testing of visual scales on-farm and at the slaughterhouse

Substantial concordance on AWI was observed among assessors (Table 2). Difficulties were found scoring some birds, mainly differentiating between low and moderate levels of CD, and between moderate and severe soiling of BS. However, our perception was not confirmed statistically, since concordance among assessors did not increase when we grouped answers (Table 2). During assessments, the inflamed skin on the breast and abdominal areas was observed to become pale within a few seconds of restraint, followed by a strong hyperaemia. No brown spot or breast blister was observed on the assessed birds.

A total of 90.7% of the birds presented moderate and severe plumage soiling and 73.4% presented CD (Figure 5). Moderate correlation was observed between litter quality and all AWI assessed on-farm, and between BS and CD (Table 2). Almost all broiler chickens were scored as poorly feathered as they presented at least one featherless area ≥ 5 cm in diameter on the breast and abdominal areas (Figure 5), therefore no correlation between feathering and the other indicators could be calculated.

Line speed varied from 5,520 to 10,080 birds per hour at the slaughterhouse because three batches were severely affected by dermatosis, which is the denomination given by the Meat Inspection Service to general skin problems in the absence of inflammation. Mean time required to assess birds was 21:23 (± 2:04) min per flock. There were difficulties assessing birds affected by dermatosis because the CS scale includes old scratches, which are characterised by lesions that resemble dermatosis in that they can present as crusts and are yellowish to brownish colour. Thus, it is advisable to ensure assessors are trained to differentiate between old scratches and other skin problems. Calculation of number of deep and superficial scratches was challenging on higher line speeds.

Discussion

Delphi methodology

The present study aimed to refine BS as a broiler chicken welfare indicator, and to develop and test two additional indicators, CD and CS. As indicated by Blokhuis *et al* (2010), involvement of stakeholders during protocol development increases its acceptability. Adherence of respondents from different groups, as well as high correlation among groups of respondents, suggest we succeeded in including relevant and knowledgeable stakeholders in this study to discuss target AWIs.

Bird-soiling is presented in the Welfare Quality® (2009) protocol as plumage cleanliness. Since birds in our study were poorly feathered, and considering that, in our experience, it is common for fast-growing broiler chickens in intensive systems to be poorly feathered, we suggested the term 'bird-soiling' to encourage assessors to assess not only the feathers, but the whole bird, including skin and feet. The use of BS instead of bird cleanliness is proposed to increase coherence between the title of the indicator and the assessment scoring system, which increases with dirtiness. When we consider the suggestion of BS and CD being scored together, it is worth noting the association between wet litter, plumage soiling and contact dermatitis (de Jong *et al* 2014). However, AWI may be more objective and consensual if they are scored separately, followed by an integration of indicators as a second step.

Our results suggest that CD has been poorly studied and therefore was not included in welfare assessment systems. The choice for the four-point scale, which included the observation of erythaema, brown spots and breast blisters, suggests that the redness observed on birds was recognised by respondents as a sign of skin irritation, and should be assessed in conjunction with other established indicators of breast lesions.

Different interpretations regarding the age of scratches in the first and second rounds probably occurred due to the notion that the animal could have experienced multiple aversive events during its life. According to Allain *et al* (2009), a broiler suffering from several lesions undoubtedly has a lower standard of welfare than one with a single lesion. Allain *et al* (2009) referred to different lesions, such as breast blisters, foot-pad dermatitis, hock burns and scratches; but it is our assumption that multiple occurrences of the same type of lesion is also indicative of a welfare problem and increased suffering.

https://doi.org/10.7120/09627286.27.3.263 Published online by Cambridge University Press

Testing of visual scales on-farm and at the slaughterhouse

Substantial concordance observed among assessors using the four-point visual scales developed in our study suggests that these are reliable for application on-farm and at the slaughterhouse. In the case of BS and CD, when scores of some categories were grouped, the number of ordinal categories was reduced to three classes. According to Nalon *et al* (2014), a scoring scale with fewer ordinal categories did not increase inter-rater reliability, and the same was observed in our study. The inclusion of a descriptive text was probably crucial in increasing concordance among assessors, by describing thresholds across levels of severity, which may not have been clear had the information been restricted merely to pictures.

The lack of variability on bird feathering scores prevented further understanding of the correlation between plumage covering and BS. Factors influencing feather growth and feather loss in broiler chickens have been studied. Dahlke *et al* (2005) suggested a negative correlation between high environmental temperature and feather coverage of fast-growing broiler chicken breeds. Poultry houses in our study had indoor temperatures higher than the recommended 18°C (Cobb-Vantress 2013), which may have contributed to general poor feathering. As early as 1978, Harris *et al* observed that poor feathering on the thigh and hip area of broiler chickens was correlated to high stocking densities, since there would be increased contact between birds, resulting in rubbing action and feather breakage. Following this rationale, greater levels of contact between the ventral body area of birds and litter may predispose to poor feathering in this area. Birds' activity may be reduced in commercial poultry houses due to high stocking densities, skeletal disorders, low lighting intensity, and the barren environment (EFSA 2010, 2012), which will lead to increased contact time between the ventral area of the body and the litter. Poor feathering increases skin exposure and, thus, is likely related to breast and abdomen skin irritation. Based on the prevalence of poorly feathered birds, we consider feathering condition as a relevant indicator to be further studied and included in broiler chicken welfare assessments. It may be an earlier indicator of welfare compared to CD. Further research is warranted to better understand the link between feathering condition and CD, and to study whether both indicators are needed and, if not, which of the two best promotes an understanding of animal welfare status.

The observed change in skin colour during bird restraint is an important item to be addressed during assessment of CD. Since handling is stressful to birds (Herborn *et al* 2015), causing an immediate rise in blood catecholamine levels (Korte *et al* 1997), the resultant peripheral vasoconstriction may have caused skin colour changes during assessment of CD. The hyperaemia observed after vasoconstriction may have been caused by reactive hyperaemia due to the accumulation of vasodilators from existing cutaneous inflammatory processes (Martinez-Lemus & Laughlin 2015). Based on this, the assessment of CD must be performed immediately after birds are restrained and prior to any other AWI. Change of skin colour did not seem to prevent the use of the scale or cause low inter-rater reliability. However, we suggest further research to establish the recovery period for normal cutaneous circulation after restraint, which may contribute to maintain best concordance between assessors for this indicator.

Moderate correlation between litter quality and all AWI measured on-farm suggested an important environmental impact on the prevalence of all types of contact dermatitis and BS. Moderate correlation between FPD and litter quality (Haslam *et al* 2007), as well the negative impact of poor litter quality on the skin of broiler chickens (de Jong *et al* 2014), have been demonstrated previously. We expected greater correlations between CD and FPD or HB, since Greene *et al* (1985) reported that contact dermatitis appears mainly on feet, followed by hock burn and breast, because hocks and breasts will increase their contact with litter as bird activity decreases (de Jong *et al* 2014). However, disagreement with regards to correlation between different types of skin lesions was also observed in other studies. For example, de Jong *et al* (2015) did not observe correlation between breast blisters and contact dermatitis; and Allain *et al* (2009) observed negative correlation between breast blisters and severe FPD and HB. Considering the early age of modern, fast-growing broiler chickens at slaughter, and depending on the litter quality, levels of FPD, HB and CD will vary and may not always be correlated. Since results concerning most correlations seem controversial, we believe it remains important to measure different animal-based indicators to assess broiler chicken welfare.

According to our data, the number of birds affected by CD was higher than birds affected by FPD and HB (Figure 5). The percentage of birds presenting CD was also higher than the mean occurrence of 0.0 to 15.8% of breast blisters and breast burns reported previously (Haslam *et al* 2007; Allain *et al* 2009; Souza *et al* 2015), which have been the only indicators considered for the health of broiler chicken breast skin. The higher percentage of skin problems reported here is a consequence of the inclusion of breast irritation, compatible with earlier signs of contact dermatitis on the breast and seems highly relevant to bird welfare assessment due to its high prevalence. We employed the term contact dermatitis for the occurrence of erythaema, which may be understood as a primary sign of contact dermatitis, especially when viewed in context with what birds were exposed to: poor abdominal feathering and the prolonged contact of skin with offending substances from excreta present on litter. According to De Jong *et al* (2014), slight redness of the breast was commonly observed in broiler chickens kept on dry litter, and skin irritation progressed towards large red areas and the presence of small brown spots as litter quality decreased (De Jong *et al* 2014). Additionally, erythaema observed in birds was considered by Delphi respondents as an unhealthy condition of the skin. However, it is our perception that histopathological studies are required to further characterise each level of macroscopic alteration described in the CD scale.

Recent studies have sought to simplify poultry welfare assessment, in an attempt to reduce assessment time and increase application of animal welfare protocols (Bassler *et al* 2013; de Jong *et al* 2015). One strategy to simplify the protocols is the correlation between assessments performed on-farm and at the slaughterhouse. Foot-pad dermatitis and hock burns have been successfully validated for fast-growing broiler chickens (de Jong *et al* 2015); and FPD has been accepted by the industry and competent authorities as a suitable indicator for identifying problems on-farm (European Commission 2017). In the case of BS, de Jong *et al* (2015) did not identify a correlation between contact dermatitis at slaughter and BS on-farm for fast-growing broiler chickens. In addition, as observed by Wilkins *et al* (2003), cleanliness of birds assessed at the slaughterhouse may be affected by pre-slaughter conditions, therefore it may not reflect litter quality or BS on-farm. Thus, we suggest that data collection on-farm remains necessary to better understand period of occurrence, prevalence and causes of certain welfare indicators, as well the correlation between animal welfare outcomes on commercial broiler chickens.

For CS assessment, modification of the carcase-sampling procedure for injuries described in the Welfare Quality® (2009) protocol allowed more detailed observation of birds, including thighs, back and both sides of carcases on the slaughter line. In addition, assessment of CS, considering different size, age and depth of scratches, may have contributed to high prevalence of this indicator. Allain *et al* (2009) already observed a high prevalence of CS (79.7 [± 13.1]%). It seems, therefore, to be an important AWI to be included during broiler chicken welfare assessment, not only because of the pain caused to birds, but also its high occurrence. High line speed was not a constraint to assess CS; however, observation of lesion depth was exhausting because of the different quantities of deep and superficial scratches allowed in each scoring level. Allain *et al* (2009) suggested that future studies should consider severity of CS, thus simplification of CS scale by not discriminating between deep and superficial scratches may not be adequate, and there is a risk of underestimating scratches. In fact, high line speed may complicate assessment when an indicator may be classified at many different levels. In contrast to de Jong *et al* (2015), who reduced levels of severity of the hock burn scale due to high line speed, our preference was to increase space between birds to be assessed. Consequently, more time was required to complete the assessment; thus, CS sampling methods require further work.

Animal welfare implications and conclusion

Results suggest that the adoption of scales for BS, CD and CS may improve our ability to assess broiler chicken welfare, since these welfare problems were prevalent and measurement consistency was acceptable. The BS scale required whole bird assessment and included pictures of birds presenting poor feathering conditions, facilitating assessment when loss of plumage is observed; a situation showing almost complete prevalence. The CD scale included hyperaemia of the breast and abdominal areas, highlighted by experts as being an unhealthy condition of the skin, which used to be overlooked and not scored during animal welfare assessments. The CS scale allowed assessment considering age, depth and length of lesions. The proposed scales for the three indicators provide both visual and descriptive information, establishing more objective thresholds between scores, which tend to increase confidence in results. Substantial concordance observed among assessors encourages application of these animal-based indicators to assess broiler chicken welfare in a wide range of poultry houses and in different countries, thereby testing the scales in a variety of animal welfare conditions.

Acknowledgements

The authors wish to thank all the professionals who contributed to test and respond to the questionnaire, partner companies, farmers, Marcos José Paulus and Nicolle Fridlund Plugge. We thank Frank Tuyttens, Elizabeth Santin and the anonymous reviewers for the valuable comments on this manuscript; and Priscilla Regina Tamioso for the English review. We also wish to acknowledge that APO Souza, VS Soriano and MA Schnaider are recipient of CAPES (Ministry of Education, Brazil) doctorate scholarships.

References

Allain V, Huonnic D, Rouina M and Michel V 2013 Prevalence of skin lesions in turkeys at slaughter. *British Poultry Science 54*: 33-41. https://doi.org/10.1080/00071668.2013.764397

Allain V, Mirabito L, Arnould C, Colas M, Le Bouquin S, Lupo C and Michel V 2009 Skin lesions in broiler chickens measured at the slaughterhouse: relationships between lesions and between their prevalence and rearing factors. *British Poultry Science 50*: 407-417. https://doi.org/10.1080/00071660903110901

Arnould C, Butterworth A and Knierim U 2009 Standardisation of clinical scoring in poultry. In: Forkman B and Keeling LJ (eds) *Assessment of Animal Welfare Measures for Layers and Broilers, Welfare Quality® Report No 9* pp 7-30. Cardiff University: Cardiff, UK

https://doi.org/10.7120/09627286.27.3.263 Published online by Cambridge University Press

Bassler AW, Arnould C, Butterworth A, Colin L, de Jong IC, Ferrante V, Ferrari P, Haslam S, Wemelsfelder F and Blokhuis HJ 2013 Potential risk factors associated with contact dermatitis, lameness, negative emotional state, and fear of humans in broiler chicken flocks. *Poultry Science 92*: 2811-2826. https://doi.org/10.3382/ps.2013-03208

Berg C 2004 Pododermatitis and hock burn in broiler chickens. In: Weeks CA and Butterworth A (eds) *Measuring and Auditing Broiler Welfare* pp 37-49. CABI Publishing: London, UK. https://doi.org/10.1079/9780851998053.0037

Blokhuis HJ, Veissier I, Miele M and Jones B 2010 The Welfare Quality® project and beyond: safeguarding farm animal well-being. *Acta Agriculturae Scandinavica 60*: 129-140. https://doi.org/10.1080/09064702.2010.523480

Cobb-Vantress 2013 *Broiler Management Guide*. http://www.cobb-vantress.com/docs/default-source/management-guides/broiler-management-guide.pdf

Dahlke F, Gonzales E, Gadelha AC, Maiorka A, Borges SA, Rosa PS, Faria Filho DE and Furlan RL 2005 Empenamento, níveis hormonais de triiodotironina e tiroxina e temperatura corporal de frangos de corte de diferentes genótipos criados em diferentes condições de temperatura. *Ciência Rural 35*: 664-670. https://doi.org/10.1590/S0103-84782005000300029. [Title translation: Feathering, triiodothyronone and thyroxine plasma levels and body temperature]

Dajani JS, Sincoff MZ and Talley WK 1979 Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change 13*: 83-90. https://doi.org/10.1016/0040-1625(79)90007-6

Dawkins MS, Donnely AE and Jones TA 2004 Chicken welfare is influenced more by housing conditions than by stocking density. *Nature 427*: 342-343. https://doi.org/10.1038/nature02226

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. *The Journal of Applied Poultry Research 23*: 51-58. https://doi.org/10.3382/japr.2013-00803

de Jong IC, Hindle VA, Butterworth A, Engel B, Ferrari P, Gunnink H, Perez Moya T, Tuyttens FAM and van Reenen CG 2015 Simplifying the Welfare Quality® assessment protocol for broiler chicken welfare. *Animal 10*: 117-127. https://doi.org/10.1017/S1751731115001706

EFSA 2010 Scientific Opinion on the influence of genetic parameters on the welfare and the resistance to stress of commercial broilers. *EFSA Journal 8*: 1-82

EFSA 2012 Scientific report updating the EFSA opinions on the welfare of broilers and broiler breeders. *EFSA Supporting Publication 9*: 116

Elfadil AA, Vaillancourt J and Meek AH 1996 Impact of stocking density, breed, and feathering on the prevalence of abdominal skin scratches in broiler chickens. *Avian Diseases 40*: 546-552. https://doi.org/10.2307/1592262

Elwinger K 1995 Broiler production under varying population densities, a field study. *Archiv fur Geflugelkunde 59*: 209-215

European Commission 2017 *Study on the application of the broilers directive (DIR 2007/43/EC) and development of welfare indicators* pp 261. European Union: Brussels, Belgium

Federici JF, Vanderhasselt R, Sans ECO, Tuyttens FAM, Souza APO and Molento CFM 2016 Assessment of broiler chicken welfare in Southern Brazil. *Brazilian Journal of Poultry Science 18*: 133-140. https://doi.org/10.1590/18069061-2015-0022

Gouveia KG, Martins da Costa P and Vaz-Pires P 2009 Welfare assessment of broilers through examination of haematomas, foot-pad dermatitis, scratches and breast blisters at processing. *Animal Welfare 18*: 43-48

Greene JA, McCracken RM and Evans RT 1985 A contact dermatitis of broilers, clinical and pathological findings. *Avian Pathology 14*: 23-38. https://doi.org/10.1080/03079458508436205

Hargis BM, Moore RW and Sams AR 1989 Toe scratches cause scabby hip syndrome lesions. *Poultry Science 68*: 1148-1149. https://doi.org/10.3382/ps.0681148

Harris GC, Mushbah M, Beasley JN and Nelson GS 1978 The development of dermatitis (scabby-hip) on the hip and thigh of broiler chickens. *Avian Diseases 22*: 122-130. https://doi.org/10.2307/1589515

Haslam SM, Knowles TG, Brown SN, Wilkins LJ, Kestin SC, Warriss PD and Nicol CJ 2007 Factors affecting the prevalence of foot pad dermatitis, hock burn and breast burn in broiler chicken. *British Poultry Science 48*: 264-275. https://doi.org/10.1080/00071660701371341

Herborn KA, Graves JL, Jerem P, Evans NP, Nager R, McCafferty DJ and McKeegan DEF 2015 Skin temperature reveals the intensity of acute stress. *Physiology & Behavior 152*: 225-230. https://doi.org/10.1016/j.physbeh.2015.09.032

Hsu C-C and Sandford BA 2007 The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation 12(10)*: 1-8

Jones BR, Facchin L and McCorquodale C 2002 Social dispersal by domestic chicks in a novel environment: reassuring properties of a familiar odourant. *Animal Behaviour 63*: 659-666. https://doi.org/10.1006/anbe.2001.1943

Jones RB and Roper TJ 1997 Olfaction in the domestic fowl: a critical review. *Physiology & Behavior 62*: 1009-1018. https://doi.org/10.1016/S0031-9384(97)00207-2

Korte SM, Beuving G, Ruesink W and Blokhuis HJ 1997 Plasma catecholamine and corticosterone levels during manual restraint in chicks from a high and low feather pecking line of laying hens. *Physiology and Behavior 62*: 437-441. https://doi.org/10.1016/S0031-9384(97)00149-2

Landis JR and Koch GG 1977 The measurement of observer agreement for categorical data. *Biometrics 33*: 159-174. https://doi.org/10.2307/2529310

Martinez-Lemus LA and Laughlin H 2015 Microcirculation, lymph and edema. In: Reece WO (ed) *Duke's Physiology of Domestic Animals* pp 376-378. Wiley-Blackwell: Iowa, USA

Muller GH 2001 Environmental skin diseases. In: Scott DW, Miller Jr WH and Griffin CE (eds) *Muller and Kirk's Small Animal Dermatology* pp 1081. Saunders: Philadelphia, USA

Nalon E, Maes D, Van Dongen S, van Riet MMJ, Janssens GPJ, Millet S and Tuyttens FAM 2014 Comparison of the inter- and intra-observer repeatability of three gait-scoring scales for sows. *Animal 8(4)*: 650-659. https://doi.org/10.1017/S1751731113002462

OIE 2013 Animal welfare and broiler chicken production systems. *OIE Terrestrial Animal Health Code*. OIE: Paris, France

Pilecco M, Almeida Paz I, Tabaldi L, Nääs I, Garcia RG, Caldara FR and Andrela G 2012 Multi-criteria analysis of the influence of rearing, equipment and catching management practices on the incidence of back scratches in broilers. *Brazilian Journal of Poultry Science 14*: 233-304. https://doi.org/10.1590/S1516-635X2012000400007

Rayens MK and Hahn EJ 2000 Building consensus using the policy delphi method. *Policy, Politics, & Nursing Practice 1*: 308-315. https://doi.org/10.1177/152715440000100409

R Core Team 2016 *R: A language and environment for statistical computing*. https://www.R-project.org

Rushen J, Butterworth A and Swanson JC 2011 Farm animal welfare assurance: science and application. *Journal of Animal Science 89*: 1219-1228. https://doi.org/10.2527/jas.2010-3589

Saraiva S, Saraiva C and Stilwell G 2016 Feather conditions and clinical scores as indicators of broilers welfare at the slaughterhouse. *Research in Veterinary Science 107*: 75-79. https://doi.org/10.1016/j.rvsc.2016.05.005

Souza APO, Sans ECO, Müller BR and Molento CFM 2015 Broiler chicken welfare assessment in GLOBALGAP certified and non-certified farms in Brazil. *Animal Welfare 24*: 45-54. https://doi.org/10.7120/09627286.24.1.045

Stephenson EL, Bezanson JM and Hall CF 1960 Factors affecting the incidence and severity of a breast blister condition in broilers. *Poultry Science 39*: 1520-1524. https://doi.org/10.3382/ps.0391520

Tuyttens FAM, Federici JF, Vanderhasselt RF, Goethals K, Duchateau L, Sans ECO and Molento CFM 2015 Assessment of welfare of Brazilian and Belgian broiler flocks using the Welfare Quality® protocol. *Poultry Science 94*: 1758-1766. https://doi.org/10.3382/ps/pev167

Veissier I, Butterworth A, Bock B and Roe E 2008 European approaches to ensure good animal welfare. *Applied Animal Behaviour Science 113*: 279-297. https://doi.org/10.1016/j.applanim.2008.01.008

Waiblinger S, Boivin X, Pedersen V, Tosi M-V, Janczak AM, Visser EK and Jones RB 2006 Assessing the human-animal relationship in farmed species: a critical review. *Applied Animal Behaviour Science 101*: 185-242. https://doi.org/10.1016/j.applanim.2006.02.001

Webster AJF 2009 The Virtuous Bicycle: a delivery vehicle for improved farm animal welfare. *Animal Welfare 18*: 141-147

Weeks CA, Nicol CJ, Sherwin CM and Kestin SC 1994 Comparison of the behaviour of broiler chickens in indoor and free-range environments. *Animal Welfare 3*: 179-192

Welfare Quality® 2009 *Welfare Quality® Assessment protocol for poultry (broilers, laying hens)*. Welfare Quality® Consortium: Lelystad, The Netherlands

Wilkins LJ, Brown SN, Phillips AJ and Warriss PD 2003 Cleanliness of broilers when they arrive at poultry processing plants. *The Veterinary Record 153(23)*: 701-703

https://doi.org/10.7120/09627286.27.3.263 Published online by Cambridge University Press