Animal (2011), **5:6**, pp 943–952 © The Animal Consortium 2011
doi:10.1017/S1751731110002594

animal

Validating indicators of sheep welfare through a consensus of expert opinion

C. J. Phythian[1†], E. Michalopoulou[1], P. H. Jones[1], A. C. Winter[1], M. J. Clarkson[1], L. A. Stubbings[2], D. Grove-White[1], P. J. Cripps[1] and J. S. Duncan[1]

[1]*Department of Epidemiology and Population Health, Institute of Global Health and Infection, Faculty of Life Sciences, University of Liverpool, Leahurst, Neston CH64 7TE, UK;* [2]*LSSC Ltd 3 Fullers Close, Aldwincle, Kettering, NN14 3UU, UK*

(Received 11 August 2010; Accepted 24 November 2010; First published online 14 January 2011)

*A consensus of expert opinion was used to provide both face and consensual validity to a list of potential indicators of sheep welfare. This approach was used as a first step in the identification of valid welfare indicators for sheep. The consensus methodology of the National Institute of Health, using pre-meeting consultation and focus group discussions, was used to ascertain the consensus opinion of a panel of sheep welfare experts. The Farm Animal Welfare Council's five freedoms were used as a framework to organise a list of current on-farm welfare issues for sheep. The five freedoms were also the welfare criterion used to identify potential on-farm welfare indicators for sheep. As a result, experts identified 193 welfare issues for sheep and lambs managed on farms across England and Wales. Subsequently, a combination of animal- (*n *= 26), resource- (*n *= 13) and management- (*n *= 22) based indicators was suggested for (i) adult rams, (ii) adult ewes (male and female sheep, over 1 year old), (iii) growing lambs (male and female sheep, over 6 weeks to 1 year old) and (iv) young lambs (male and female lambs, 6 weeks old and under). The results from this study could therefore be used to inform the further development of valid methods of assessing the on-farm welfare of sheep.*

Keywords: sheep, animal welfare, indicator, expert opinion, validity

Implications

Valid welfare indicators are required nationally and internationally by farmers, veterinary surgeons, the sheep industry and farm inspectors in order to measure and monitor on-farm welfare standards for benchmarking, certification, farm assurance and legal purposes. Currently, on-farm welfare inspections rely on the assessment of management- and resource-based measures. However, there are increasing demands for measures that assess the quality of an animal's life using direct animal-based outcomes. Compared to other farm animal species, few validated animal-based measures of sheep welfare currently exist. Therefore, this study has the potential to inform and enhance future on-farm assessments of sheep welfare.

Introduction

Consumer awareness and concern for farm animal welfare has provided the impetus for driving the provision of products from high welfare systems, assurance of farm animal welfare standards and demands for welfare labelling of food (Waiblinger *et al*., 2001; Blokhuis *et al*., 2003; Main *et al*., 2003b). In addition to non-regulatory purposes, welfare assessments are undertaken to assess compliance with legislative and regulatory standards (EFSA, 2007). Consequently, there is a need for valid, reliable and feasible measures for use in on-farm welfare assessment systems which would be transparent and fair to both producers and consumers (FAWC, 2009; Mullan *et al*., 2009).

In the past, on-farm welfare assessments focused on the evaluation of farm resources (e.g. provision of housing and feed) known as resource-based indicators or those that assess farmer policies and management practices (management-based indicators; Capdeville and Veissier, 2001; FAWC, 2009; Mullan *et al*., 2009). Since animal welfare can be considered to be the outcome of the interaction between genotype, management and the environment (Sørensen and Fraser, 2010), a valid assessment of welfare should also include animal-based welfare outcomes (FAWC, 2009; Mullan *et al*., 2009; Sørensen and Fraser, 2010).

A valid welfare indicator would be a genuine measure of animal welfare (EFSA, 2007; Bracke *et al*., 2008). Since

† E-mail: C.J.Phythian@liverpool.ac.uk

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

welfare measures (or indicators) are akin to diagnostic tests, they may be evaluated in a similar manner. Essentially, a new diagnostic test would be conferred as valid, if it produced the same result as an established reference test (Greiner and Gardner, 2000). However, this approach is problematic for animal welfare assessment as a reference test does not currently exist (de Passillé and Rushen, 2005). An alternative approach is to use a consensus of expert opinion to judge the validity of welfare indicators that are incorporated into on-farm assessment protocols. By judging the suitability and relevance of each measure, expert opinion provides consensual and face validity to the selection of welfare indicators (Abramson, 1974; Johnsen *et al*., 2001; Scott *et al*., 2001; Bracke *et al*., 2002). Previously, expert opinion has identified a number of welfare indicators for cattle, poultry and pigs (Capdeville and Veissier, 2001; Whay *et al*., 2003a; Botreau *et al*., 2009) and has also been used to identify a number of welfare issues for sheep (Cronin *et al*., 2002; Waterhouse *et al*., 2003; Pines *et al*., 2007; Phillips *et al*., 2009).

The methods for gathering the consensus of expert opinion vary, from using Delphi questionnaires (Whay *et al*., 2003a; Phillips and Phillips, 2010), online surveys (Fernie *et al*., 2008), small group meetings such as the Nominal Group Technique (NGT; Delbecq *et al*., 1975) to conference-style meetings (Glaser, 1980; NIH, 1990; Pines *et al*., 2007; Rousing *et al*., 2007). The NIH consensus development programme (NIH, 1990) is a well-recognised conference-style method that has been used in human medicine to identify numerous clinical and biotechnological issues (www.consensus.nih.gov). This method is relevant for developmental research studies as experts meet face-to-face enabling idea generation and open discussions (NIH, 1990). The NIH approach can involve a pre-meeting consultation, such as the use of questionnaires to give experts greater time to enlarge on the topic of interest (Glaser, 1980). In addition, experts may be split into sub-groups during the conference to allow an in-depth discussion of all the relevant issues (Glaser, 1980).

The overall aim of this study was to establish valid, reliable and feasible indicators for the on-farm assessment of sheep welfare. This paper describes the initial identification of a number of sheep welfare indicators which encompassed the welfare concepts in the Farm Animal Welfare Council (FAWC) five freedoms (FAWC, 1994) and were sensitive to the current on-farm welfare issues for sheep. An NIH approach (NIH, 1990) allowed a panel of experts to judge the validity (face and consensual) of sheep welfare indicators that were included in on-farm assessment studies. Following this initial validation step, the internal validity (evaluation of bias, reliability and diagnostic accuracy) and the external validity (generalisability, feasibility and applicability) of these indicators would be evaluated in field studies (Greiner and Gardner, 2000). Thereafter, the ranking or exclusion of any indicators could be based on the evidence of their diagnostic validity, reliability, and feasibility for use in on-farm welfare applications.

Material and methods

Expert panel selection

The objective of the expert panel selection process was to identify a broad panel of expert members (Garabed *et al*., 2009). An 'expert' was defined as a person having a minimum of 10 years sole experience of sheep farming in the UK sheep industry and/or professional achievements in industries and organisations allied to sheep farming, veterinary services and welfare research. Experts were short-listed by members of the research team to avoid overlap in areas of representation (Garabed *et al*., 2009). Following approval by the Department for the Environment, Farming and Rural Affairs (Defra), 33 experts were formally invited to attend a 1 day expert panel meeting.

Before their attendance, experts were given clear, written guidance on the objectives of the meeting and the required output of the expert panel – to produce a list of indicators of sheep welfare, which encompassed all aspects of the five freedoms and were sensitive to current on-farm welfare issues for sheep in England and Wales. In addition, experts were informed of their role at the meeting, that is, to validate the identification of potential indicators of sheep welfare. Where invited members did not respond to follow-up contact ($n = 2$), or were unable to attend ($n = 3$), a substitute member with similar expertise was selected from the short-list and invited to attend. As a result, 33 members confirmed their attendance.

Identifying on-farm welfare issues for sheep – postal worksheet

One month before the meeting, a one-page worksheet was distributed to all experts (electronic and postal hard copy) in order to identify potential on-farm welfare issues for the individual sheep. Each expert was asked to list a maximum of 10 issues considered to affect the on-farm welfare of (i) ewes (female sheep over 1 year old), (ii) rams (male sheep over 1 year old), (iii) growing and fat lambs (any sex, over 6 weeks to 1 year old) and (iv) young lambs (from birth to 6 weeks old) in England and Wales. Experts were asked to consider the factors that affected the welfare at the level of the individual sheep. Anonymous responses from completed worksheets were reviewed by the authors who categorised the suggested welfare issues into the relevant area(s) of the five freedoms framework (FAWC, 1994). The five freedoms detailed by the FAWC are accepted as a sound framework for the development of animal welfare indicators (Whay *et al*., 2003a; FAWC, 2009). This information was summarised into a document that was distributed in electronic and paper versions to all panel members 1 week before the expert panel meeting (Figure 1).

Identifying on-farm welfare issues for sheep – expert panel meeting

The expert panel meeting was held on 23 February 2008, and consisted of 30 members from welfare organisations and charities ($n = 2$), sheep welfare research ($n = 5$), welfare

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

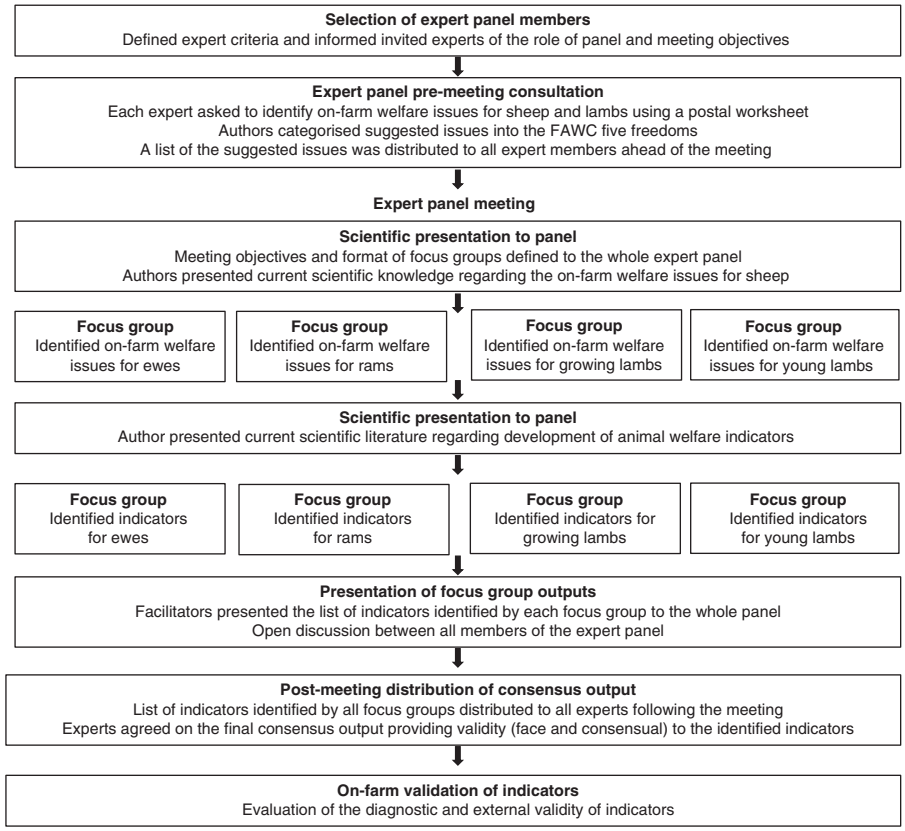


Figure 1 Flow-chart of the expert consensus of opinion method. Structure of the consensus of opinion methodology summarising the steps involved in the validation of the sheep welfare indicators. The final step in this flow-chart summarises the subsequent validation stages, in which the diagnostic and external validity of indicators are evaluated in on-farm studies.

inspectors ($n = 1$), veterinary animal welfare advisors ($n = 2$), government advisors on animal welfare policies ($n = 2$), veterinary sheep specialists ($n = 3$), veterinary surgeons from general practice ($n = 4$), sheep farmers ($n = 5$), sheep industry services ($n = 2$), and sheep production consultants ($n = 4$). The meeting followed NIH guidelines (NIH, 1990), in which an introductory presentation on the objectives of the research project was given and members were informed of their specific role in the validation of welfare indicators. The expert panel comprised of members with different experiences and expertise in sheep health, welfare and production. Therefore, in line with NIH meetings (NIH, 1990), the authors gave the expert panel a presentation on the current, scientific knowledge of on-farm welfare issues for sheep (Figure 1), identified through a systematic literature search (C. Phythian, 2008, unpublished). For the purposes of this study, welfare was defined as a factor, event or action that affected one or more of the five freedoms of an individual sheep (FAWC, 1994). The panel was then divided into four focus groups (Figure 1). Each group focused on identifying the welfare issues for one of four production stages: (i) ewes, (ii) rams, (iii) growing and fat lambs and (iv) young lambs. The aim was to allow a balance of opinions within each group and a mix of members from the different representative areas was allocated into each group. A focus group comprised of seven or eight experts, plus a facilitator, who was a trained member of the project team (Glaser, 1980; Krueger and Casey, 2009). In addition, two non-participating recorders maintained accurate written minutes of group discussions. The welfare issues document, which categorised all the welfare issues identified from the pre-meeting worksheet into the five freedoms framework (FAWC, 1994), was distributed to the panel 1 week ahead of the expert panel meeting. This provided the experts with more time to consider the welfare issues. At the meeting, this document was used as the basis for focus group discussions. All panel members were given the opportunity to consider, discuss, modify and clarify the welfare issues listed in this document and were free to alter the categorisation of any welfare issue within the five freedoms framework. As well as guiding the identification of welfare indicators by the experts, this document was also one of the final outputs of the expert panel meeting – akin to the consensus of expert opinion statement produced in NIH conferences (NIH, 1990).

Identifying potential welfare indicators – expert panel meeting

The next stage in the expert panel process was to identify valid indicators of sheep welfare based on the suggested welfare issues for sheep and the five freedoms framework (FAWC, 1994). Following the NIH process (NIH, 1990), the authors gave a presentation to the panel on the current

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

scientific knowledge of animal welfare indicators, including those developed for other species. Experts then returned to their focus groups (Figure 1) and were asked to consider and identify measures capable of assessing each of the issues listed within the five freedoms framework. Experts were asked to identify indicators that were animal-based (where possible), non-invasive and practicable for assessments performed under farm conditions. Once the focus groups had completed this task, the panel reconvened and each facilitator gave a presentation on the valid indicators identified by their group. All members of the panel were then asked to comment on the output of the other focus groups. Following the guidance of NIH consensus methods, an official output for the consensus of expert opinion was produced by the authors after the meeting (NIH, 1990). This consisted of a final document listing the on-farm welfare issues and indicators for sheep as suggested by the expert panel. This document was electronically distributed to all experts within 1 week of the meeting. Experts were then requested to read the document and confirm that the indicators and issues were those as suggested at the meeting. If the majority of experts agreed with the meeting outputs, then a consensus of opinion would be established (NIH, 1990).

Results

On-farm welfare issues for sheep

A total of 19 out of 33 worksheet responses was received (response rate of 58%). Experts suggested 193 potential welfare issues for sheep in England and Wales. Of these, 53 on-farm welfare concerns were identified for the ewe production stage. A total of 45 issues for rams, 53 issues for growing and fat lambs and 42 issues for young lambs were also raised. Table 1 lists the welfare issues identified in the postal worksheet by at least 50% of experts ($n \geqslant 17$). The pre-meeting consultation found that lameness was consistently identified as an on-farm welfare issue for all four of the production stages. For young lambs, mutilation procedures of tail docking and castration were considered to be key welfare issues (Table 1). Poor body condition was i dentified as an on-farm welfare issue for both ewes and rams, whereas gastro-intestinal parasitism of growing lambs and dystocia in ewes were issues specific to these production stages.

At the panel meeting, experts identified and categorised a total of 193 welfare issues for ewes, rams and lambs into the relevant criteria of the five freedoms framework (FAWC, 1994). Since all the experts indicated their agreement with the post-meeting output documents, a consensus of expert opinion on the welfare issues for sheep was achieved.

Given the large number of issues identified by the panel, a summary of the welfare issues identified for each freedom criteria was produced (Table 2). The authors grouped and summarised the 193 issues into a list of general welfare issues, which covered all four production stages ($n = 58$). Where appropriate, issues specific to the welfare of ewes, rams, young and growing lambs were highlighted. This organisation provided 14 general welfare issues within the freedom from hunger and thirst, 14 within the freedom from discomfort, 11 within the freedom from pain, injury and disease, 8 for freedom to express normal behaviour and 10 general welfare issues were produced under the freedom from fear and distress criterion. For example, within the freedom from pain, injury or disease, the expert panel suggested that a large number of diseases and health conditions were on-farm welfare issues for sheep and lambs. Accordingly, specific diseases and health conditions were grouped and summarised by the authors, for example, contagious ovine digital dermatitis, foot rot and scald were summarised as 'infectious foot lesions' (shown in brackets in Table 2). Subsequently, the general disease conditions were grouped into an overall, general welfare issue, which was termed 'heath status' (Table 2).

On-farm welfare indicators for sheep

The animal-, resource- and management-based indicators suggested by the panel are shown in Table 3. Of the animal-based indicators suggested, 10 were identified for ewes, 13 for rams, 11 for young lambs and 9 for growing and fat lambs. Many of the welfare indicators identified for each production stage were similar. Since there was overlap across the focus group outputs, the authors grouped all the suggested measures into a single list of animal- ($n = 26$), resource- ($n = 13$) and management-based ($n = 22$) indicators (Table 3). Although most indicators could be used to assess all four production stages, the assessment of time taken to stand and suck, along with an overall indicator of lamb thrift was specifically suggested for young lambs. Resource- and management-based indicators suggested by the four focus groups also showed similarity, so these indicators were also grouped into a single list (Table 3). Resource-based measures included the assessment of housing and

Table 1 *Welfare issues identified through a pre-meeting postal worksheet*

Young lambs	Growing lambs	Ewes	Rams
Castration (63%)	Gastrointestinal parasitism (74%)	Poor body condition, poor nutrition (84%)	Poor body condition, poor nutrition (68%)
Tail docking (58%)	Lameness (53%)	Lameness (84%)	Lameness (63%)
Lameness, septic arthritis (53%)	Myiasis, maggot control (53%)	Dystocia (58%)	–[§]

Perceived welfare issues identified by pre-meeting consultation by more than 50% of respondents ($n = 17$) are displayed together. The percentage of respondents who identified the issue to be of importance for each of the sheep production stages is shown in brackets.
[§]No additional issues suggested by more than 50% of respondents.

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

Table 2 *On-farm welfare issues for sheep and lambs*

Freedom criteria	Welfare issues
Freedom from hunger and thirst	Provision of appropriate and adequate feed and forage
	Provision of a continuous supply of clean water
	Appropriate body condition for production stage and purpose
	Colostrum management (immunoglobulin concentration, adequate volume)[YL]
	Management and nutrition of orphan lambs (artificial rearing methods)[YL]
	Neonatal lamb behaviour (time taken to stand and suck)[YL]
	Maternal behaviour (ewe-lamb bonding)[YL]
	Culling policy
	Management of dietary change (bought-in animals, weaning[YL, GL])
	Selection of animals suited to management system (able to thrive on grass)
	Management flexible to changing animal requirements (parturient ewe nutrition[E])
	Appropriate stocking density
	General health status
	Dental health (condition of molar and incisor teeth)
Freedom from discomfort	Provision of shelter and shade for grazed animals
	Provision of lie-back area for grazed animals (grazing root crops)
	Clean, dry environmental and pasture conditions
	Appropriate fleece cover for turnout post-shearing
	Use of appropriate breed or genotype (ability to withstand climatic conditions)
	Stocking density for housed and grazed animals
	Provision and quality of bedding in housing
	Flooring surface and hygiene
	Management of hypothermia
	Provision of appropriate winter housing
	Provision of exercise for housed animals
	Year-round care and inspection
	Fitting and maintenance of equipment (correct fitting of harness and raddle[R])
	Appropriate rearing for intended management and purpose
Freedom from pain, injury or disease	Year round inspection (cast sheep)
	Appropriate intervention at lambing (management of dystocia)[E, YL]
	Transport of pregnant ewes and lame sheep
	Presence of horns (physical injuries and blindness due to horn occlusion)
	Appropriate body conformation (double muscling)
	Predation[YL]
	Reproductive management (mis-mating)[R, E, GL]
	Quality of equipment and buildings (design and maintenance of housing and handling facilities, correct fitting of raddle and harness[R])
	Management practices associated with pain (use of electro-ejaculator[R], vasectomy[R], lambing intervention[E, YL], ear tagging and notching, castration and tail docking[YL, GL])
	Health status (presence of specific clinical signs and diseases including lameness, joint disease, infectious foot lesions, nutrient deficiencies and toxicities, metabolic disease, skin lesions, reproductive disease, inheritable defects, endo- and ecto-parasitism, dental health, mastitis, infectious disease, ocular conditions, respiratory disease)
	Disease prevention and control (vaccination, anthelmintic control, appropriate action and treatment of disease, stockperson inspection, veterinary input, culling policy, bio-security measures)
Freedom to express normal behaviour	Appropriate intervention at lambing[E, YL]
	Reproductive management practices (use of artificial insemination)
	Space allowance: opportunity for exercise for indoor-housed animals
	Provision of environmental enrichment for housed animals
	Group size and composition (overcrowding, mixing unfamiliar sheep)
	Isolation of individual sheep (sick animals)
	Artificial rearing of lambs [YL]
	Maternal behaviour (ewe-lamb bonding, mis-mothering)[E, YL]
Freedom from fear and distress	Appropriate intervention at lambing[E, YL]
	Artificial insemination
	Intensive finishing systems[GL]
	Quality of handling: stockperson skills and quality of handling facilities
	Method and timing of weaning

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

Table 2 *Continued*

Freedom criteria	Welfare issues
	Use of lamb adopter (tethering of ewes[E])
	Transport of pregnant ewes and lame sheep
	Group size and composition (stocking rates, ewe to ram ratios)
	Ability of humans to recognise and interpret sheep behaviour
	Use and control of dogs (dog worrying, use of aggressive farm dogs)

[R] Welfare issue specific to rams.
[E] Welfare issue specific to ewes.
[GL] Welfare issue specific to growing lambs.
[YL] Welfare issue specific to young lambs.

Table 3 *Animal-, resource- and management-based indicators of sheep welfare identified by a consensus of expert opinion*

Animal-based indicators	Resource-based indicators	Management-based indicators
Body condition	Water provision, access and quality	Mortality and culling records
Live weight	Quality and quantity of food	Abattoir feedback
General demeanour	Trough space	Medicine records
Alertness to approach in the field	Space allowance in housing	Disease records
Atypical behaviours	Flooring type	Movement records
Separation from the flock	Provision and quality of bedding	Faecal egg count records
Play behaviours	Presence of an isolation area	Growth rate records
Ruminatory behaviour	Good fencing and farm boundaries	Scanning records
Qualitative behavioural assessment	Sward height	Policy for hypothermia management
Panting	Stocking density at grazing	Ear tagging policy and quality
Cleanliness (rear and belly)	Provision of shelter and shade	Observation of specific management procedures (tagging, castration)
Time spent standing	Presence of handling facilities	Observation of daily farmer tasks
Gait/lameness	Presence of a lie back area	Assessment of handling skills
Skin condition		Castration and tail- docking policy
Fleece condition and wool cover		Stockperson awareness of on-farm disease
Urolithiasis		Judgement of appropriate on-farm action
Tail length		Current flock health plan
Eye condition		Mixing of horned and un-horned sheep
Nasal discharge		Mixing of different ages and sizes
Ear integrity (tagging injuries)		Presence of carcasses with live sheep
Biochemical measures		Farmer self-assessment
Rumen fill		Reproductive management and policy
Presence of in-growing horns		
Body injuries and wounds		
Time to stand and suck[YL]		
Thriving lamb assessment[YL]		

[YL] Indicators specific to young lambs.

associated facilities and provisions, for example, space allowance and quality of bedding. The panel identified that castration and tail docking procedures were painful procedures (Tables 1 and 2). However, experts suggested that it may be more feasible to ascertain whether these mutilations were practised, the method used, and the on-farm policy regarding analgesia and anaesthesia use, rather than measuring the pain responses of lambs (Molony and Kent, 1997).

Experts were also asked to consider the feasibility of any suggested measures (Table 3). Assessments under working farm conditions were considered to pose several difficulties, including the location and gathering of extensive flocks, handling of pregnant ewes and gathering of ewes with lambs at foot. Experts therefore suggested that the nature of the sheep production year should be considered before assessment visits were undertaken. Following distribution of the post-meeting document, two panel members made additional comments, which were duly noted and recorded. These comments did not affect the list of identified indicators (Table 3), but they helped to inform the on-farm evaluation studies. Experts confirmed the validity of the indicators listed in the post-meeting consensus document. Therefore, a consensus of expert opinion provided validity to the selection of welfare indicators to undergo on-farm testing on sheep in England and Wales.

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

Discussion

Expert validation of welfare indicators

In the absence of a reference test (or gold standard) for welfare assessment (de Passillé and Rushen, 2005), previous welfare research has also based the selection and initial validation of welfare indicators on expert opinion (Cronin *et al.*, 2002; Whay *et al.*, 2003a; Bracke *et al.*, 2008). This is because expert opinion is considered to provide both face and consensual validity to the welfare indicators (Abramson, 1974; Johnsen *et al.*, 2001; Scott *et al.*, 2001). The study presented here is the first to determine valid sheep welfare indicators for the on-farm welfare assessment of sheep using an NIH consensus approach (NIH, 1990). The NIH methodology provided validation of welfare indicators that were identified by expert opinion. The consensus method appears to be suitable for other animal health and welfare research applications, for example, agreeing on standard laboratory techniques or defining a specific animal welfare condition. Before applying this method, there are several factors that need to be considered. First, the definition of who is an 'expert' and what is their level of expertise (Spoolder *et al.*, 2003). Second, expert opinion can differ according to occupation (Bracke *et al.*, 2008; Fernie *et al.*, 2008). Therefore, a biased view of welfare may be ascertained by using an unbalanced panel with experts from specific occupations (Vang, 1986). Furthermore, consensus methodologies make assumptions on the quality of the expert panel's decisions (Fink *et al.*, 1984). Expert opinion may have little in common with widely known scientific facts or reasoning (Vang, 1986) and specialists may concentrate on a small and specific issue instead of considering the whole picture (Vang, 1986). Expert opinion should alter as scientific knowledge advances, although experts can ignore welfare research findings and provide opinion based solely on their personal experiences (Main *et al.*, 2003a). In view of these issues and in the absence of a reference test for animal welfare, a combined approach using a scientific literature review (C. Phythian, 2008 unpublished) and expert opinion guided the identification of valid welfare indicators in this study (Figure 1). The advantage of employing expert opinion, over the opinion of the study researchers, is that a group of experts would be expected to possess greater diversity in experience and knowledge of the subject under discussion, which is beneficial to an idea-generation process. As a number of concerns are recognised with the use of expert opinion, these factors were accounted for in both study design and conduct. In particular, the following considerations were made: (i) clear definition of experts (Garabed *et al.*, 2009), (ii) selection of experts according to level of expertise, (iii) balanced composition of members in terms of expertise in sheep welfare and occupation, (iv) quality of decisions, by using a sound consensus method (NIH, 1990), (v) clear communication of meeting objectives and fulfilment of distinct tasks (Kynn, 2008) and (vi) presentation of scientific knowledge to establish an educated and informed panel.

Methods for gathering a consensus of opinion

The Delphi method has been suggested as a feasible way of collecting expert opinion by avoiding conflicts between individual opinions and thereby minimising bias of opinion (Delbecq *et al.*, 1975; Whay *et al.*, 2003a). However, Delphi questionnaires can be criticised on the grounds of poor response rates and long response times (Hsu and Sanford, 2007). Furthermore, postal questionnaire methods are unsuitable for research requiring personal and direct communications (Fink *et al.*, 1984). Here, face-to-face Delphi questionnaire methods, recently used to identify individual farmer concerns for sheep welfare (Phillips and Phillips, 2010) may prove useful. Individual interviews were not considered to be appropriate for this study, instead, a method that elicited the opinion of a number of experts during a 1-day meeting and encouraged discussion and ideas-generation was required. As a result, other face-to-face consensus methods including the NGT (Delbecq *et al.*, 1975), Improved Nominal Group Technique (Fox, 1989) and the NIH method (NIH, 1990) were examined. The selected NIH method allows researchers to present the most current, scientific knowledge and concepts to experts attending a consensus-style meeting (NIH, 1990). NIH can include a pre-meeting consultation to prevent a 'bottleneck' in group discussions and also to reduce the meeting time required (Fox, 1989). Focus groups, as used in NIH meetings, allow for idea generation and exchange of expertise and opinions (Glaser, 1980; Fitzpatrick and Boulton, 1994; Krueger and Casey, 2009). In this study, each focus group was comprised of seven to eight members, guided by trained facilitators (Glaser, 1980) to encourage full contribution and minimise over-domination by any member (Delbecq *et al.*, 1975; Krueger and Casey, 2009). To reduce opinion biases, groups were mixed in terms of occupation, experience and expertise, and this also prevented the group from deferring to the opinion of any one individual (Kynn, 2008; Krueger and Casey, 2009). There is a lack of a definition on the establishment of a consensus in decision-making processes (Scott and Black, 1991). The NIH method, takes the 'majority rule' approach (NIH, 1990) to the identification of consensus, so, one disadvantage is that minority or extreme views may be ignored. The goal of this study was to identify broad expert agreement, so it may be appropriate to discount very polarised or minority opinions (Scott and Black, 1991). The final step in NIH methods allows experts to comment and deliberate on the scientific evidence before finally offering their judgement (NIH, 1990). The post-meeting distribution of the output of this study (Figure 1) permitted individual experts to review the welfare indicators further and confirm to the authors their agreement of the list of welfare indicators as created at the expert panel meeting. As a result, this approach provided both face and consensual validity to the sheep welfare indicators identified by the expert panel (Table 3).

Welfare issues identified

This study identified a number of on-farm welfare issues affecting sheep. Although experts were asked to specifically identify issues that were relevant to farms in England and

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

Wales, many of the welfare issues raised are applicable to other countries, systems and stages of production, such as transport and lairage (Cronin *et al.*, 2002; Waterhouse *et al.*, 2003; Phillips and Phillips, 2010). Furthermore, to the author's knowledge, this is the first study to validate potential welfare indicators that consider: (i) sheep of all ages (neonates to adult sheep), (ii) all stages of on-farm production (birth, lactation, weaning, growth, reproduction) and (iii) the diversity in sheep management (intensive, indoor-managed flocks to large extensive systems). The pre-meeting worksheet identified thin ewes (84% of respondents), hypothermia of young lambs and endo-parasitism of growing lambs (74%) were on-farm welfare issues. Pre-meeting consultation also found that both lameness and poor body condition were consistently identified as sheep welfare issues (Table 1). Interestingly, the welfare issues suggested by this expert panel concur with the expert opinion regarding the welfare issues for extensively managed sheep (Waterhouse *et al.*, 2003), and for sheep managed in Australia (Cronin *et al.*, 2002; Phillips *et al.*, 2009; Phillips and Phillips, 2010). In contrast to other studies (Cronin *et al.*, 2002; Rousing *et al.*, 2007; Phillips *et al.*, 2009; Phillips and Phillips, 2010), the suggested welfare issues were not ranked in terms of their importance. Ranking can be used to create an overall welfare index (Fernie *et al.*, 2008), or to identify a set number of measures that are used to assess on-farm welfare standards (Rousing *et al.*, 2007). However, the creation of an overall welfare index or overall welfare assessment system for sheep was not the objective of this study. Instead, the aim was to identify a number of valid welfare measures, essentially individual diagnostic tests, which would undergo further validation in on-farm studies. A fundamental aspect of this study was to view welfare from the animal's point of view by encompassing both physical and mental elements as detailed in the FAWC five freedoms (1994). As the authors consider that each of the five freedoms is important to the welfare of the individual sheep, ranking was not appropriate for this initial study stage. Indeed, ranking of measures at this stage risked omitting a welfare issue that was specific to a certain time point or to a particular farm management system. The authors intended that any ranking, synthesis or exclusion of the suggested indicators would be based on the results of diagnostic validity, reliability and feasibility studies (Figure 1).

Potential sheep welfare indicators

Before this study, indicators for the on-farm assessment of sheep in England and Wales had not been identified. Previous work has identified potential animal-based welfare indicators for sheep transported by sea (Pines *et al.*, 2007) and for organic sheep managed in Italy (Napolitano *et al.*, 2009). Animal-based indicators have received increasing attention for their value in on-farm welfare assessments (Whay *et al.*, 2003b; Knierim and Winckler, 2009; Mullan *et al.*, 2009). A combination of animal-, resource- and management-based indicators may be most appropriate for on-farm welfare assessments (Capdeville and Veissier, 2001; Waiblinger *et al.*, 2001). The panel therefore generated a range of animal-based, resource-based and management-based indicators that they considered to be practical given on-farm assessment conditions and the limits of a 1-day assessment period (Table 3).

Animal-based indicators. Many of the animal-based indicators suggested by this expert panel were focused on measures of flock health, for example, body condition scoring (BCS), lameness scoring, and skin lesion assessment. Although BCS was identified as a useful welfare assessment tool, and has been scientifically validated as a measure of the body fat content of sheep (Russel *et al.*, 1969), the panel recognised that BCS was not an inappropriate indicator for young and growing animals. The panel also suggested that fewer categories of condition scores may be more appropriate for welfare assessment purposes. For example, sheep could be categorised as simply 'fit', 'fat' or 'thin'. Another specific indicator suggested by the panel included the use of an overall young lamb thriving score in order to categorise lambs as either 'thriving' or 'ill-thriven'. This composite indicator was designed by the panel members to provide an overall welfare score based on the assessment of a number of health and welfare measures such as posture, demeanour, standing ability and also specific health-based measures such as abdominal fill, gait, body condition and the presence of any ocular abnormalities. Experts also identified the need to assess animals managed at specific times during the sheep production calendar. They also emphasised the importance of including artificially reared lambs in any young lamb welfare assessments.

Experts suggested that an assessment of sheep behaviour was a useful welfare assessment tool and that behavioural separation of individual animals from the flock and/or the expression of a dull, depressed demeanour clearly indicated a health and/or welfare issue for the individual sheep. In addition, the use of a whole-animal behavioural approach to sheep welfare assessment was proposed. The expert panel identified that qualitative behavioural assessment; a holistic approach to the assessment of animal behavioural expression (Wemelsfelder and Lawrence, 2001), should also be included in the list of animal-based indicators (Table 3).

Overall, the panel's suggestions closely followed those of another expert panel (Waterhouse *et al.*, 2003) who proposed that on-farm issues for sheep welfare may be considered in terms of the severity, duration and intensity of the effect on welfare. In this project, the intensity was considered to be the number of sheep affected by a specific indicator. There is a lack of definition between individual and flock welfare indicators (Waterhouse *et al.*, 2003), but in this study, experts were asked to consider welfare issues that were important for the individual sheep. Although it was clear from the outputs of each focus group that experts identified issues and indicators that were relevant to the welfare of the individual sheep, they suggested that the interpretation of indicator assessments could be based on a flock-level approach. Indeed, the panel in this study suggested that the flock prevalence of indicators was the most valid and feasible way of performing on-farm assessments of

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

the animal-based indicators. Experts were informed that any indicators initially suggested by the panel would undergo further validation in the next stage of the project, when diagnostic test evaluation of the indicators would be performed in field studies. This approach permitted the study of both individual and group-level methods of assessing sheep. Results of on-farm testing of indicators would be presented to the expert panel at subsequent meetings in order to allow experts to make an informed and science-based decision on the interpretation of individual- and flock-based welfare assessments. The panel also suggested that, where appropriate, the severity and duration of the welfare indicators could be considered in any scoring systems that were developed by the authors.

As this is the first step in identifying valid indicators of sheep welfare, not all the animal-based indicators suggested by the panel may be applied during field testing. Experts suggested that biochemical measures (Table 3), such as plasma cortisol, could be used for on-farm assessment of sheep welfare, although they also suggested that measures may not be feasible for a whole flock assessment. Interestingly, experts did not suggest any non-invasive biochemical measures such as faecal cortisol levels. Accordingly, the authors did not incorporate any biochemical measures into the field testing of indicators.

Resource- and management-based indicators. The indicators suggested by the panel (Table 3) broadly fell into (i) assessments of grazing, housing and on-farm facilities, (ii) assessments of farm records, and, finally, (iii) discussions with farmers to ascertain on-farm management practices. The effect of season and production system should be considered in the application of resource-based indicators, for example, sheep may only be housed during short and specific periods and often receive intermittent supplementary feeding. For a large proportion of the sheep production calendar, the main resource for sheep in the United Kingdom may be considered to be the provision of grazing. Therefore, resource-based assessments would be likely to be focused on the quality and the quantity (sward height) provided (Table 3).

Conclusion

The objective of this study was to identify valid indicators for the on-farm assessment of sheep welfare. This study followed good practice guidelines regarding meeting design, conduct of focus groups, and selection of experts in order to use a consensus of expert opinion to validate potential indicators of sheep welfare. Although the NIH approach is widely used in human medicine, it has not previously been reported for use in expert consultation processes for animal health and welfare research purposes. A clear advantage for expert consultations that require direct and face-to-face communication of a number of members is the balanced, multi-disciplinary, science-based and informed approach offered by the NIH technique. Experts identified a large number of on-farm welfare concerns for sheep and also suggested a range of potential animal-, resource- and management-based indicators. The use of the method employed here may be appropriate for the development of welfare indicators for other species. This expert validation process was the first step in the identification of valid indicators of sheep welfare. The next step is to investigate the diagnostic validity (accuracy and reliability) and feasibility of these measures during field studies.

Acknowledgements

The authors gratefully acknowledge the role of Defra in funding this project (AW1025, Development of indicators for the on-farm assessment of sheep welfare) and thank all members of the expert panel for their contribution and support with the project.

References

Abramson JH 1974. Validity. In Survey methods in community medicine, pp. 151–164. Churchill Livingstone, New York.

Blokhuis HJ, Jones RB, Geers R, Miele M and Veissier I 2003. Measuring and monitoring animal welfare: transparency in the food product quality chain. Animal Welfare 12, 445–455.

Botreau R, Veissier I and Perny P 2009. Overall assessment of animal welfare: strategy adopted in welfare quality. Animal Welfare 18, 363–370.

Bracke MBM, Metz JHM, Spruijt BM and Schouten WGP 2002. Decision support system for overall welfare assessment in pregnant sows B: validation by expert opinion. Journal of Animal Science 80, 1835–1845.

Bracke MBM, Edwards SA, Engel B, Buist WG and Algers B 2008. Expert opinion as 'validation' of risk assessment applied to calf welfare. Acta Veterinaria Scandinavica 50, 29–40.

Capdeville J and Veissier I 2001. A method of assessing welfare in loose housed dairy cows at farm level, focusing on animal observations. Acta Agriculturae Scandinavica Section A – Animal Science 51, 62–68.

Cronin GM, Barnett JL, Edge JK and Hemsworth PH 2002. Identifying animal welfare issues for sheep in Australia. International Journal of Sheep and Wool Science 50, 534–550.

de Passillé AM and Rushen J 2005. Can we measure human-animal interactions in on-farm animal welfare assessment? Some unresolved issues. Applied Animal Behaviour Science 92, 193–209.

Delbecq AL, Van de Ven AH and Gustafson DH 1975. Group techniques for program planning. In A guide to nominal group and Delphi processes (ed. AC Filley) Scott, Foresman and Company, Illinois.

EFSA (European Food Safety Authority) 2007. Basic information for the development of the animal welfare risk assessment guidelines. (ed. S Berbieri and C Nassuato), Parma 2007. EFSA/AHAW/2006/01, pp. 1–36. Retrieved 1 November 2007, from http://www.efsa.europa.eu/en/scdocs/doc/130ar.pdf.

FAWC (Farm Animal Welfare Council) 1994. Report on the welfare of sheep: PB 1755. FAWC, London.

FAWC (Farm Animal Welfare Council) 2009. Report on farm animal welfare in Great Britain: past, present and future. FAWC, London.

Fernie AC, Tribe A, Murray PJ and Phillips CJC 2008. The creation and implementation of a Great Ape welfare index. In Proceedings of the International Primatological Society XXII Congress (ed. P Lee, H Buchanan-Smith, P Honess, A MacLarnon and B Sellers). Edinburgh, Scotland, 361pp.

Fink A, Kosecoff J, Chassin M and Brook RH 1984. Consensus methods – characteristics and guidelines for use. American Journal of Public Health 74, 979–983.

Fitzpatrick R and Boulton M 1994. Qualitative methods for assessing health care. Quality in Health Care 3, 107–113.

Fox WM 1989. The improved nominal group technique. Journal of Management Development 8, 20–27.

Garabed RB, Perez AM, Johnson WO and Thurmond MC 2009. Use of expert opinion for animal disease decisions: an example of foot-and-mouth disease status designation. Preventive Veterinary Medicine 92, 20–30.

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press

Glaser EM 1980. Using behavioural science strategies for defining the state-of-the-art. Journal of Applied and Behavioural Science 16, 79–92.

Greiner M and Gardner IA 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. Preventive Veterinary Medicine 45, 3–22.

Hsu C and Sanford BA 2007. The Delphi technique: making sense of consensus. Practical assessment, research and evaluation 12. Retrieved 12 April 2010, from http://www.joe.org/joe/2010february/rb2.php.

Johnsen PF, Johannesson T and Sandøe P 2001. Assessment of farm animal welfare at herd level: many goals, many methods. Acta Agriculturae Scandinavica Section A- Animal Science 51, 26–33.

Knierim U and Winckler C 2009. On-farm welfare assessment in cattle: validity, reliability and feasibility issues and future perspectives with special regard to the welfare quality approach. Animal Welfare 18, 451–458.

Krueger RA and Casey MA 2009. Focus groups: a practical guide for applied research, 2nd edition. SAGE Publishing, Los Angeles.

Kynn M 2008. The 'heuristics and biases' bias in expert elicitation. Journal of the Royal Statistical Society Series a-Statistics in Society 171, 239–264.

Main DCJ, Whay HR, Green LE and Webster AJF 2003a. Preliminary investigation into the use of expert opinion to compare the overall welfare of dairy cattle farms in different farm assurance schemes. Animal Welfare 12, 565–569.

Main DCJ, Kent JP, Wemelsfelder F, Ofner E and Tuyttens FAM 2003b. Applications for methods of on-farm welfare assessment. Animal Welfare 12, 523–528.

Molony V and Kent JE 1997. Assessment of acute pain in farm animals using behavioral and physiological measurements. Journal of Animal Science 75, 266–272.

Mullan S, Edwards SA, Butterworth A, Whay HR and Main DCJ 2009. Interdependence of welfare outcome measures and potential confounding factors on finishing pig farms. Applied Animal Behaviour Science 121, 25–31.

Napolitano F, De Rosa G, Ferrante V, Grasso F and Braghieri A 2009. Monitoring the welfare of sheep in organic and conventional farms using an ANI 35 L derived method. Small Ruminant Research 83, 49–57.

NIH (National Institute of Health) 1990. Consensus development at the NIH: improving the programme. Report of a study by a committee of the Institute of Medicine council on health care technology. National Academy Press, Washington DC. Retrieved 4 January 2008, from http://www.nap.edu/catalog.php?record_id=1563.

Phillips CJC and Phillips AP 2010. Attitudes of Australian sheep farmers to animal welfare. Journal of International Farm Management 5th edition 2, 1–26.

Phillips CJC, Wojciechowska J, Meng J and Cross N 2009. Perceptions of the importance of different welfare issues in livestock production. Animal 3, 1152–1166.

Pines MK, Petherick JC, Gaughan JB and Phillips CJC 2007. Stakeholders' assessment of welfare indicators for sheep and cattle exported by sea from Australia. Animal Welfare 16, 489–498.

Russel AJF, Doney JM and Gunn RG 1969. Subjective assessment of body fat in live sheep. Journal of Agricultural Science 72, 451–454.

Rousing T, Jakobsen IA, Hindhede J, Klaas IC, Bonde M and Sørensen JT 2007. Evaluation of a welfare indicator protocol for assessing animal welfare in AMS herds: researcher, production advisor and veterinary practitioner opinion. Animal Welfare 16, 213–216.

Scott E and Black N 1991. When does consensus exist in expert panels? Journal of Public Health Medicine 13, 344–344.

Scott EM, Nolan AM and Fitzpatrick JL 2001. Conceptual and methodological issues related to welfare assessment: a framework for measurement. Acta Agriculturae Scandinavica Section A – Animal Science 51, 5–10.

Sørensen JT and Fraser D 2010. On-farm welfare assessment for regulatory purposes: issues and possible solutions. Livestock Science 131, 1–7.

Spoolder H, De Rosa G, Horning B, Waiblinger S and Wemelsfelder F 2003. Integrating parameters to assess on-farm welfare. Animal Welfare 12, 529–534.

Vang J 1986. The consensus development conference and the European experience. International Journal of Technology Assessment in Health Care 2, 65–76.

Waiblinger S, Knierim U and Winckler C 2001. The development of an epidemiologically based on-farm welfare assessment system for use with dairy cows. Acta Agriculturae Scandinavica Section A – Animal Science 51, 73–77.

Waterhouse T, Colgrove P, Vipond J and Goddard A 2003. Identification of key issues – a summary of the final group session. In Improving sheep welfare on extensively managed flocks: economics, husbandry and welfare (ed. PE Goddard), pp. 73–77. Scottish Agricultural College, Aberdeen, Scotland.

Wemelsfelder F and Lawrence AB 2001. Qualitative assessment of animal behaviour as an on-farm welfare-monitoring tool. Acta Agriculturae Scandinavica Section A – Animal Science 51, 21–25.

Whay HR, Main DCJ, Green LE and Webster AJF 2003a. Animal-based measures for the assessment of welfare state of dairy cattle, pigs and laying hens: consensus of expert opinion. Animal Welfare 12, 205–217.

Whay HR, Main DCJ, Green LE and Webster AJF 2003b. Assessment of the welfare of dairy cattle using animal-based measurements: direct observations and investigation of farm records. Veterinary Record 153, 197–202.

https://doi.org/10.1017/S1751731110002594 Published online by Cambridge University Press