Characterization of hazards, welfare promoters and animal-based measuresfor the welfare assessment of dairy cows: Elicitation of expert opinion

Luigi Bertocchi[a], Francesca Fusi[a,*], Alessandra Angelucci[a], Luca Bolzoni[b], Stefano Pongolini[b],

Rosa Maria Strano[a], Jessica Ginestreti[a], Giorgia Riuzzi[a], Paolo Moroni[c,d], Valentina Lorenzi[a]

[a] *Italian National Reference Centre for Animal Welfare, Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna "Bruno Ubertini", Via A. Bianchi 9,25124 Brescia, Italy*

[b] *Risk Analysis Unit, Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna, Sezione di Parma, Via dei Mercati 13/A, 43100 Parma, Italy*

[c] *Università degli Studi di Milano, Dipartimento di Medicina Veterinaria, Via Celoria 10, 20133 Milan, Italy*

[d] *Animal Health Diagnostic Center, Quality Milk Production Services, Cornell University, Ithaca, NY 14853, United States*

Keywords:

Dairy cattle

Loose housing system

On-farm animal welfare assessmentNon-animal-based measure

Animal-based measure

A B S T R A C T

An expert opinion elicitation, based on a modified Delphi technique, was organized to collect the opinion of 16 Italian veterinarians with the aim of conducting a hazard and a welfare promoter characterization for defining and weighing a list of management and housing factors potentially associated with negative or positive welfare outcomes in dairy cows kept in loose housing systems. In addition, the 16 experts judged a set of animal-based measures in order to rate them by appropriateness and by the level of animal pain and suffering due to the welfare consequences they measure. Veterinary experts were asked to score 52 hazards, 47 welfare promoters and 18 animal-based measures. Management and housing hazards, that were determined to be associated with a very high impact on the welfare of dairy cows, were mainly referred to lactating cows (absence of bedding material, presence of inadequate or slippery floor in walking areas, wrong design of the lying area), but also the use of harmful tools for animal handling and the lack of scheduled foot inspection, trimming and foot bathing reached very high impact scores. Management and housing welfare promoters dealing with optimal cow comfort around resting obtained the highest ratings, together with the presence of experienced and trained stockpersons, the implementation of an attentive animal grouping strategy and the control of temperature, humidity and ventilation. Considering animal-based measures, the highest ranking of importance was given to observation of lameness, records of mortality of adult cows and calves, observation of integument alterations and body condition scoring. This study has been the starting point for the development of the first Italian national protocol for the assessment of the welfare of dairy cows farmed in loose housing systems.

1. Introduction

At European Union level there are currently no specific rules for dairy cow welfare (EFSA, 2012b, Section "Background as provided bythe European Commission", p.6) and the Council Directive 98/58/CE (European Union Council, 1998) for the protection of animals kept for farming purposes is considered general, vague and non-efficient in protecting animals throughout Europe (Broom, 2017). In addition, public concern over livestock production is rising and most of the Europeans would like to receive more information about the conditions under which farmed animals are kept (European Commission, 2016).To fill these gaps and to efficiently safeguard animal welfare at farm level, researchers in this field are currently trying to develop assessment methods that are practical but also scientifically valid (Walker et al., 2014). In the last years, research on animal welfare assessment has been focused mainly on the use of animal-based measures (ABMs), which are directly related to the animal's experience and its ability to cope with the given environment (de Vries et al., 2013). ABMs were chosen by the European Welfare Quality project (WQ) as the best tools to assess the actual welfare state of farmed animals and to identify the most criticaland urgent welfare problems (Welfare Quality, 2009). However, WQ protocols cannot be applied with high frequency at farm level, because they are too time-consuming (Heath et al., 2014). In fact, ABMs are assumed to be more valid than resource and management-based indicators (also referred to as non-animal-based measures – N-ABMs), but their use could make farm audit schemes complicated, less feasible, and costly (de Vries et al., 2013; Andreasen et al., 2014; Grandin, 2014). In several cases, N-ABMs could be more efficient than ABMs andthey could help to identify hazards potentially associated with the ob- served welfare outcomes (EFSA, 2012b, Section 2.3.2, p.32; Lundmarket al., 2016). Both ABMs and N-ABMs are necessary in an assessment protocol for achieving an effective overall classification of animal welfare at farm level. The choice of the measures and the most ap- propriate mix between ABMs and N-ABMs depends on the aim of the welfare assessment (EFSA, 2012b, Section "Conclusions and re- commendations", p.44).

To date, in Italy there are not protocols including both ABMs and N- ABMs that can be used routinely at farm level to assess the welfare ofthe over 1,800,000 dairy cows present within the national territory (ISTAT, 2017). Thus, in order to support the official controls, to pro-mote the implementation of on-farm assessment of dairy cow welfare and to provide consumers with information, the Italian National Re- ference Centre for Animal Welfare (CReNBA) carried out a study aimedat developing a feasible and simplified on-farm animal welfare assess- ment protocol for dairy cows in loose housing systems. This study triedto fill the gap between research and on-farm application and to define public standards that go beyond the minimum legislative requirements, for actually improving dairy cow welfare at national level and for overcoming the recently outlined problems linked to the proliferationof private standards (More et al., 2017).

Guidelines for the risk assessment for animal welfare have been published by the European Food Safety Authority (EFSA), adapting the well-established risk assessment methodologies for animal diseases and food safety (EFSA, 2012a). The proposed methodology can be used to assess not only factors that could be associated with negative welfare outcomes, but also factors that could positively influence the welfare of the animals (EFSA, 2012a, Section 3.4, p.15). Following these guide- lines, CReNBA carried out an expert opinion elicitation in order to conduct both a hazard characterization (HC) and a welfare promoter characterization (WPC), as a first step of a "risk-benefit assessment" for the welfare of dairy cows in loose housing systems. The aim was to characterize and weigh a set of management and housing factors and to include them as N-ABMs in a welfare assessment protocol that will be used at national level. In addition, the experts were asked to weigh a list of ABMs, which will be integrated in the same protocol, in order to assess the actual animal welfare state, too. To our knowledge, this is the first time that data from a WPC for dairy cows farmed in loose housing systems are published.

2. Materials and methods

In this study, semi-quantitative HC and WPC applied to dairy cow welfare and based on an expert opinion elicitation were carried out following the EFSA "Guidance on risk assessment for animal welfare"(EFSA, 2012a) and the EFSA scientific opinions on the welfare of dairy cows and sheep (EFSA, 2009a, 2014a). In addition, a set of ABMs was submitted to the judgment of the experts involved in the elicitation process. The opinion of each expert was gathered using a modified Delphi technique (Lees and Lievaart, 2013). The Delphi process is a very useful method for collecting, in a formalized and transparent way, the opinions and the judgments of experts and practitioners on a par- ticular subject (Yousuf, 2007). In this survey, a modified Delphi process was used, that means the first step of this method was omitted, because the list of factors (hazards and welfare promoters) and of measures (ABMs), that had to be judged by the experts, were previously prepared by a different group of veterinarians and not provided directly by the experts (as in Lees and Lievaart, 2013). Thus, pre-selected factors and measures were submitted to the experts for collecting their opinion and the obtained data were analyzed (first round of the expert opinion). Then, the experts received back their questionnaires (together with the average scores of the group) and they had the chance to modify their answers (second round of the expert opinion) (EFSA, 2012b, Section 2,p.11). According to the Delphi process, this second round was carried out in order to reach a better consensus within the group of experts (EFSA, 2012b, Section 2.4.4, p.37).

1. *Identification of the target population*

Dairy cows farmed in loose housing systems (both cubicle houses and straw yards) were the target population identified for the survey. This choice was justified by the fact that these are the most common farming systems for dairy cows in Italy. All the animal life stages withinan ordinary dairy farm were taken into account: calves (animals from birth to 6 months of age), heifers (young non-productive animals, from6 months of age to the 7th month of pregnancy), lactating cows (con- sidering an average lactation period of 305 days) and dry cows (con- sidering an average dry period of 60 days and including heifers between the 8th month of pregnancy and the calving).

2. *Selection of the factors and of the welfare consequences*

The selection of the factors, to be submitted to the experts for the HC and the WPC, was previously carried out by a team of 10 Italianveterinarians, during a focus group. This interview technique generally involves 8–12 participants, who discuss a clear topic under the direc-tion of a moderator (Stewart and Shamdasani, 1990). The participants

were CReNBA's collaborators and advisors, who were involved in 12 sessions of interview at CReNBA's headquarters, during a whole year.The credentials of the members of the focus group are reported in

Supplementary Table S1.

During the focus group, the list of recommendations published by EFSA (2012b, Appendix A, pp.51–69) was used as a guideline for theselection of the factors and the related N-ABMs. In fact, this list can be considered as a "toolbox" from which picking out the most appropriate measures according to the aim of the animal welfare assessment (EFSA, 2012b, Section 2.1.2, p.25). Measures able to identify the fulfillment of EFSA's high priority recommendations for dairy cow welfare (EFSA, 2012b, Appendix A, pp.51–69) were preferred during the selection process together with measures able to verify the compliance with the requirements of the Italian laws (D. Lgs. N. 146/2001; D. Lgs. N. 126/ 2011). Only factors related to dairy cow management and housing were taken into account. Genetic selection was not considered since its ef-fects are lifelong and they cannot be managed in the short period (Smulders, 2009). During the focus group, for each selected manage- ment and housing factor, 3 different intensities of animal exposure were defined, based on scientific publications (e.g. EFSA, 2009e; Welfare Quality, 2009) and national laws. If no data could be extracted from these sources, national databases and individual expertise were dis- cussed and used. Starting from an intensity of exposure considered not to influence dairy cow welfare (baseline level), a critical level and a benefit level were established and specified for each factor. Accordingto EFSA (2012a, Appendix A, p.19), the baseline level corresponds to an intensity of exposure that could be associated with animal welfare outcomes usually recognized as standard in an average husbandry system; the critical level (hazard) is an intensity of exposure that couldbe potentially associated with a state of harm; while, the benefit level corresponds to an intensity of exposure that could make a factor a potential welfare promoter. The chosen management and housing fac- tors and the corresponding intensity specifications are listed in Sup- plementary Table S2. In total, 52 factors were selected by the focus group to be proposed to experts during the elicitation process. These factors could potentially influence the welfare of either the entire herd or a specific group of animals (i.e. calves, heifers, lactating cows or dry cows). For eachfactor, a hazard level and a baseline level were identified, while a benefit level was specified only for 47 factors, since for 5 factors (i.e. type of handling, milking parlor access and exit, use of concentrate feeds, bedding material for newborn calves in single pens and possibi- lity for calves to see and touch each other) no evidence of welfare en- hancement beyond the baseline level was available from literature orfield experience.

During the focus group, 5 macro categories of welfare consequences on the target population were also identified. These categories were chosen according to EFSA scientific opinions on dairy cow welfare (EFSA, 2009a, 2009b, 2009c, 2009d) and considering welfare criteriaand measures included in the WQ assessment protocol for dairy cows (Welfare Quality, 2009). The following categories were recognized of great importance for dairy cow welfare: i) udder health, ii) metabolic needs, iii) locomotion and foot health, iv) integument alterations, v) behavior. Each of the 52 selected factors was assumed to be potentially associated with one or more, negative or positive, consequences within these 5 welfare categories and the possible link between each factor and each group of effects was investigated during the expert opinion elici- tation.

3. *Selection of the animal-based measures*

In addition to HC and WPC, indicators for the on-farm measuring of welfare outcomes were also rated by the experts involved in the elici- tation process. Consequences on animal welfare can be measured by one or more welfare indicators, which could be direct (e.g. body con- dition score, observation of integument alterations, observation of la-meness, etc.) or indirect ABMs (e.g. bulk milk somatic cell count, on-farm mortality, etc.) (Welfare Quality, 2009; EFSA, 2012a, Section "Glossary and abbreviations", p.28; EFSA, 2012b, Section 2.1.2, p.24).

The list of measures to be submitted to the expert opinion elicitation was drawn up during the previously mentioned focus group. The se-lection was based on scientific evidence (Welfare Quality, 2009; EFSA, 2012b, Appendix B, pp.70–73) and by taking into account that these measures should be routinely recorded in Italian dairy farms. Where already available, valid, reliable and feasible ABMs were preferred during the selection process. ABMs were, in fact, chosen from the "toolbox" provided by EFSA (2012b, Appendix B, pp.70–73), which also included WQ's ABMs. All the WQ's measures had been reported to possess face validity and many of them can be considered sufficiently valid, reliable and feasible (Knierim and Winckler, 2009). In addition, some of the ABMs listed by EFSA (2012b, Appendix B, pp.70–73), and not included in the WQ protocol, have been already tested for validity, reliability and feasibility (EFSA, 2012b, Section 2.1.1 and Appendix B, pp.13 and 70–73). Only ABMs able to measure negative welfare con- sequences were taken into account, because current scientific knowl- edge does not offer valid, robust and feasible indicators for directly and routinely measuring positive welfare outcomes (Boissy et al., 2007; Yeates and Main, 2008). Usually, ABMs are recorded on-farm at individual level, but once measured they could be aggregated at herd or group level (e.g. per- centage of animals with integument alterations) to interpret them against defined threshold values (EFSA 2012b, Section 1.3, p.10). The value at which animal welfare can be considered acceptable or un- acceptable depends on the aim of the welfare assessment (EFSA, 2012b, Summary, p.4). For this reason, during the focus group, for each selected ABM, 3 gradual threshold values (i.e. ABM levels 1–3) were also defined, based on the available scientific literature (e.g. Welfare Quality, 2009; Nielsen et al., 2014) or on national databases and in- dividual experience. The 18 selected ABMs, the group to which they are referred (i.e calves, heifers, lactating cows, dry cows, all groups), the measured adverse effects and the 3 defined levels are listed in Sup- plementary Table S3.

4. *Selection of the experts*

For the expert opinion elicitation, a public competition was organized for the setup of the expert panel. The activity that was going to be asked, the time involved and the possibility of travel reimburse- ment were described in the announcement for the experts to be aware. The required profile had to fulfill the following general selection cri- teria: degree in veterinary medicine, at least 10 years of experience in dairy cows (i.e. physiology, health, behavior, welfare, etc.) and/or dairy cow farming, and availability for travelling and taking part to a 2-day technical meeting at CReNBA's headquarters. A resume in Europass format was required. The public competition was publicized on CReNBA's website and disseminated by e-mails to the Italian public veterinary services and to the members of the biggest Italian veterinary associations for dairy cattle.

Forty-six candidates answered to the announcement and sent their resume, but 4 were rejected because they did not satisfy the selection criteria. The candidates were mainly public veterinarians and practi- tioners, but also 2 academics and 2 farmers presented their candidacy. Due to resource availability, 16 was the maximum number of experts that could be hosted, reimbursed and managed for this activity. EFSA (2014b, Section A.2.4.2, p.159) reported that 8–15 experts could be a viable number when workshop or face-to-face approaches are used for the expert opinion elicitation. In fact, larger groups do not change the findings significantly and they could incur in extra expenses and extra time (Aspinall, 2010).

An initial group of 22 veterinarians was selected and ranked based on the received resumes. The selection process was carried out by taking into account the level of experience and also with the aim of guaranteeing expert heterogeneity in competences and in geographical working areas. Following the established ranking order, the experts were contacted by telephone and the final group of 16 veterinarians was set up. The expert panel, engaged for the elicitation exercise, was composed by 3 females and 13 males, average age was 49.9 years (range 39–58 years old). The group included 16 veterinarians with different skills: 6 public veterinarians with experience in dairy cow welfare, ethology, health, management and housing; 2 academics, skilled in dairy cow welfare, nutrition and production; 2 farmers graduated in veterinary medicine and currently engaged in dairy farming and 6 practitioners (e.g. nutritionist, clinician, farm consultant, etc.) (Supplementary Table S4). For the purpose of the study, they were invited to take part in a 2-day technical meeting to familiarize with the elicitation process and to acquire the basic concepts of risk assessment, HC and WPC. During the training, experts were given explanations and detailed instructions about the task. They were asked to fill out three questionnaires individually at home, without consulting each other and using their personal experiences, their scientific or technical knowledge and the published data (Hultgren et al., 2016). The time given to complete the task was 1 month.

5. *Hazard and welfare promoter characterizations*

The 16 experts were asked to weigh the 52 factors selected by the focus group. The factors were divided in hazards (52) and welfare promoters (47), based on the corresponding intensity specifications (hazard level or benefit level, respectively). As a consequence, 2 sepa- rate semi-quantitative characterizations, one for the hazards and the other one for the welfare promoters, were carried out. During the2 days of training, the factors and the corresponding intensity specifi- cations (i.e. baseline, hazard and benefit levels) were accurately ex- plained to the experts, by means of pictures, videos and examples, in order to allow standardized estimates. The experts were asked to weigh the potential negative or positive impact of each proposed factor on the welfare of dairy cows in relation to the 5 identified welfare categories(i.e. udder health; metabolic needs; locomotion and foot health; in- tegument alterations and behavior). As previously described, the factors could be referred to different animal groups (i.e. calves, heifers, lac- tating cows, dry cows) or to the entire herd (animal group = all) and separate scoring was required for each group. Experts were considered to have sufficient knowledge about dairy cow exposure to the factorsand about the consequences of the exposure, so all the necessary data were collected using experts' knowledge elicitation (Hultgren et al., 2016).

In the HC, experts were asked to estimate the magnitude (scoring scale from 0 – none – to 3 – high) and the likelihood (from 0% to 100%) of the negative welfare consequences, if any, that could be associatedwith the exposure of the animals to each of the 52 given hazards. If a hazard was considered potentially associated with different adverse effects, impairing more than 1 of the 5 identified welfare categories, separated scores were assigned to each category (Ribó and Serratosa, 2009). To score the magnitude, experts were asked to give an overall value, integrating the severity of the adverse effects, their duration and how often they are repeated during the life time of the animal, as in EFSA (2014a, Section 2.1.4, p.16). For each given hazard, if the sug- gested intensity of exposure was not considered by the experts to re- present a potential critical level and no associated adverse effects could be recognized in 1 or more of the 5 identified welfare categories, they were asked to assign a magnitude score of 0 (none) at the corresponding category. Considering the likelihood of the adverse effects, this para- meter was defined as the most likely percentage of animals, in a typical Italian loose housing dairy farm, that will experience the adverse effects following the exposure to the given hazard. This definition was adapted from EFSA (2014a, Section 2.1.4, p.15) and the reference period was setto 1 year. Finally, experts were asked to rate their certainty in relationto the likelihood value they provided (EFSA, 2014a, Section 2.1.4, p.16). Certainty was scored as "high", "medium" and "low" (as in EFSA, 2014a, Section 3.1.4, p.42), according to Table 1.

WPC was carried out using the same methodology developed for HC, but taking into account the potential animal welfare improvements associated with the exposure of dairy cows to the welfare promoters. Expert opinion was elicited to score 47 welfare promoters with the potential to positively influence 1 or more of the 5 identified welfare categories. Welfare promoters were characterized by the intensities for a total of 780 entries; while, for welfare promoters 705 ratings were asked (15 values for each of the 47 factors). The task was carried out by the experts individually at home and the spreadsheets were sent back to CReNBA's staff by e-mail.

6. *Animal-based measure weighing*

The ABMs selected by the focus group were scored by the experts involved in the elicitation process. During the 2-day technical meeting, the selected measures and their on-farm application were explained to the experts. Individual exercises were carried out by assessing pictures and videos in order to test experts' comprehension of the measures and to allow standardized estimates. Reference material (short explanatory manual with pictures, captions and references) was left to the experts to support their task and they were invited to contact CReNBA's staff incase of any technical doubts during the fulfillment of their work.

Firstly, experts were asked to score each of the selected ABMs in relation to their appropriateness and to the magnitude of the animalwelfare impairment caused by the adverse effects they measure. According to EFSA (2012c, Section 2.1, p.13), the appropriateness of an ABM refers to both its feasibility and its capability of correctly assessinga specific welfare consequence, and it was judged by the expert using a scoring scale from 1 (low) to 3 (high). The magnitudes of the negative welfare consequences assessed by the ABMs were also scored (scoring scale from 1 – low – to 5 – high). For this latter rating, experts wereasked to weigh the amount of pain and suffering of the affected animals,

taking into account both the severity and the duration of the welfare consequences. Then, level 1 of each ABM, deemed to represent the acceptable threshold value in the Italian dairy cow population, was scored 0 by default and expert knowledge was elicited to weigh the severity of the welfare impairment of the entire herd or of a specific group of animals at levels 2 and 3 (scoring scale from 0 – negligible – to 3 – severe). The 16 experts were also asked to rate the certainty of these latter answers according to Table 1. provided in the questionnaire (see intensity specification of the benefit For ABMs weighing, a MS-Excel spreadsheet (Microsoft level in Supplementary Table S2). Magnitude of positive welfare con-sequences was scored on a 4–point scale (from 0 – none – to 3 – high) similar to the one used in HC, but considering different degrees of welfare improvement. For each welfare promoter, if the proposed intensity of exposure was not considered by the experts to represent a potential benefit level and no associated positive welfare consequences could be recognized in at least 1 of the 5 proposed welfare categories, they were asked to assign a magnitude score of 0 (none) at the corre- sponding category. Likelihoods of positive welfare consequences were rated on a scale from 0% to 100% and certainty levels (Table 1) were assigned to these opinions. Separate scores were required for each re- cognized association between welfare promoters and welfare con- sequences in a specific category.

In total, experts were asked to fill out 2 MS-Excel spreadsheets (Microsoft Corporation, Redmond, WA, USA), one for the 52 hazards and the other one for the 47 welfare promoters. For each hazard 15 quantities were required (3 values for each of the 5 welfare categories), Table 1 Certainty rates, modified from Smulders (2009) and Hultgren et al. (2016). The overall impact score of each welfare promoter (*i*), $IS(WP_i)_{overall}$, was obtained using the same method applied for the hazards. First of where $magnitude(H_{i,j,k})$ represents the magnitude of the welfare con- sequences potentially associated with hazard *i* in the welfare category *j* and provided by expert *k*; $likelihood(H_{i,j,k})$ represents the percentage of all, the impact score of the *i*-th welfare promoter on the *j*-th welfarecategory and provided by the expert *k* (namely $IS(WP_{i,j,k})$) was com- puted as follows:

Corporation, Redmond, WA, USA) with 108 entries (6 values for each of the 18 ABMs) was submitted to each expert. The questionnaire was filled out individually at home and all the opinions were sent back to CReNBA's staff by e-mail.

7. *Data analysis*

All the experts sent back the 3 filled out questionnaires. Answerswere checked for typing errors and data were analyzed to calculate the average scores assigned by the group of experts. According to the Delphi process, after 1 week, experts received back the average scores and their individual excel files and they had the chance to modify their answers and to correct typing errors. Errors were corrected and 12 experts (75.0%) modified their answers, adjusting the scores either upwards or downwards. For the analysis of the raw data, obtained from the expert opinion elicitation, the methods proposed by EFSA (EFSA, 2009a, Section 1.1.5,p.12; EFSA, 2012a, Box 3, p.15 – therein –; EFSA, 2014a, Section 3.1.4, p.42) were used with the necessary adjustment.

1. *Hazard impact score calculation* valuation Explanation Based on the experts' ratings, an overall impact score, namely *IS* (H_i)$_{overall}$, was determined for each hazard (*i*). First of all, the values igh Solid and complete data available with strong evidence provided in many references and similar conclusions reported by the majority of the authors; otherwise, considerable and consistent field experience Medium Some or no complete data available with evidence provided in fewreferences and differences in authors' conclusions; otherwise, limited evidence from field experience Low Scarce or no data available with evidence provided in unpublished reports and great differences between authors' conclusions; otherwise, scarce field observations given by the experts were standardized between 0 and 1 (as in EFSA, 2014a, Section 3.1.4, p.42). Ordinal values of 0, 0.33, 0.67 and 1 were assigned to the 4 magnitude classes and likelihood ratings (0%–100%) were converted to a score between 0 and 1. According to EFSA (2014a, Section 3.1.4, p.42), the 3 certainty classes, i.e. "high", "medium" and "low", were converted in uncertainty scores of 0.25, 0.5 and 1, re- spectively. As first step, the impact score of the *i*-th hazard on the *j*-th welfare category assigned by the expert *k* (namely $IS(H_{i,j,k})$) was
2.
3. Fig. 1. Overall impact scores of management hazards ($IS(H_i)_{overall}$) and welfare promoters ($IS(WP_i)_{overall}$).
4. All = all animal groups; DC = dry cows; HF = heifers;LC = lactating cows.

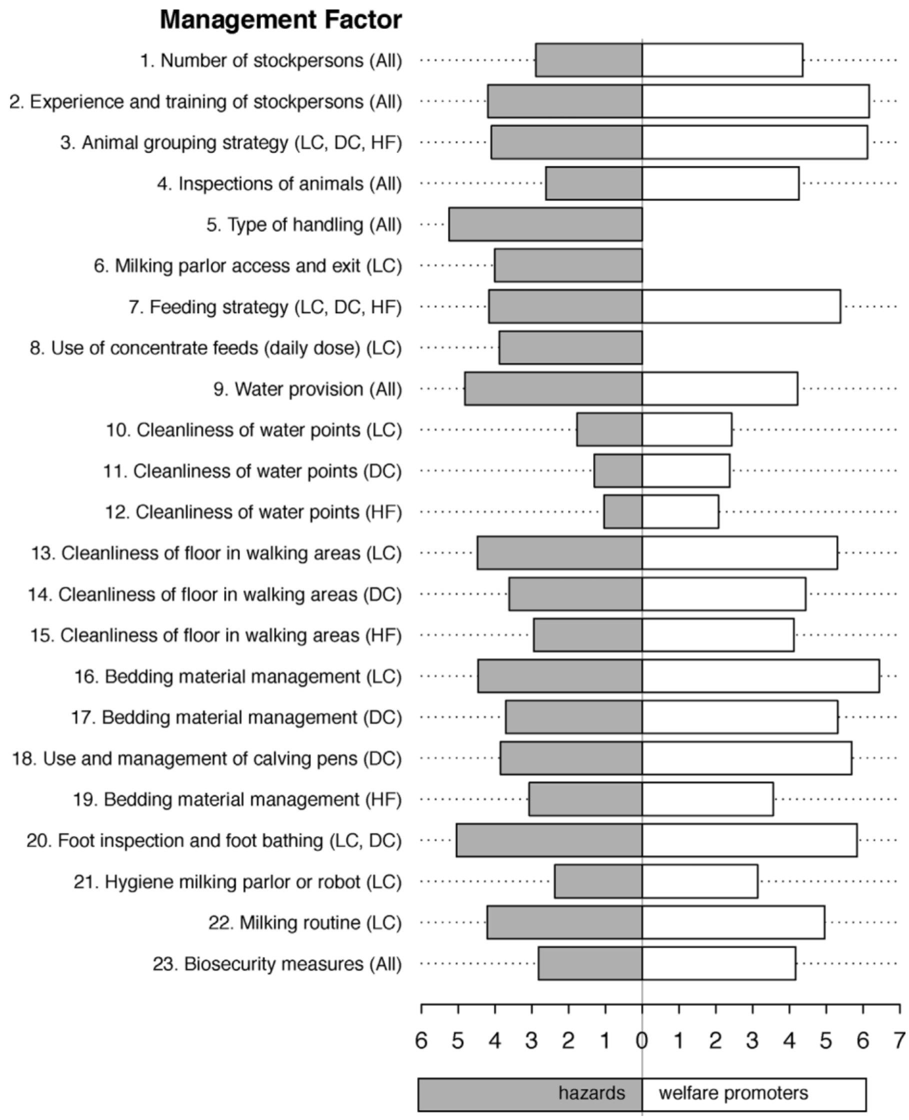

Fig. 1. Overall impact scores of management hazards ($IS(H_i)_{overall}$) and welfare promoters ($IS(WP_i)_{overall}$).

All = all animal groups; DC = dry cows; HF = heifers; LC = lactating cows.

computed as follows: *2.7.2. Welfare promoter impact score calculation* $IS(H_{i,j,k}) = magnitude(H_{i,j,k}) \times likelihood(H_{i,j,k})/uncertainty(H_{i,j,k})$ (1)

animals affected by the consequences potentially associated with hazard *i* in the welfare category *j* and provided by expert *k*; *uncertainty* $IS(WP_{,j,k}) = magnitude(WP_{i,j,k}) \times likelihood(WP_{i,j,k})/uncertainty(WP_{i,j,k})^{(H}i,j,k)$ represents the uncertainty rating provided by expert *k* and referred to the likelihood of the consequences associated with hazard *i* in the welfare category *j*.

Then, category impact scores, namely $IS(H_{i,j})$, referred to the effect of the *i*-th hazard on the *j*-th welfare category, were obtained by com- puting the mean of the $IS(H_{i,j,k})$ provided by the 16 experts:

$IS(H_{i,j}) = mean\,[IS(H_{i,j,k})]$. According to EFSA (2014a, Section 3.1.4, *k* p.42), the uncertainty-weighted average was chosen as data aggregation method. If a hazard was found to be potentially associated with negative welfare outcomes within different welfare categories, the amount of suffering resulting from the different adverse effects was assumed to be additive (EFSA, 2012a, Appendix A, p.25). Therefore an overall impact score for each hazard *i* was calculated by summing up the obtained category impact scores: $IS(H_i)_{overall} = \sum_j IS(H_{i,j})$. where $magnitude(WP_{i,j,k})$ represents the magnitude of the consequences potentially associated with welfare promoter *i* in the welfare category *j* and provided by expert *k*; $likelihood(WP_{i,j,k})$ represents the percentage of animals, which could benefit from the consequences associated with welfare promoter *i* in the welfare category *j*, provided by expert *k*; $uncertainty(WP_{i,j,k})$ represents the uncertainty rating provided by expert *k* and referred to the likelihood of the consequences associated with welfare promoter *i* in the welfare category *j*. As for the hazards, cate- gory impact scores of welfare promoter *i* on the welfare category *j*, $IS(WP_{i,j})$, were computed as the average value of the $IS(WP_{i,j,k})$ provided by the 16 experts: $IS(WP_{i,j}) = mean\,[IS(WP_{i,j,k})]$. Individual category *k* scores, $IS(WP_{i,j})$, were then aggregated to determine the overall impact score of each welfare promoter *i*: $IS(WP_i)_{overall} = \sum_j IS(WP_{i,j})$.

2.7.3. Animal-based measure impact score calculation

An overall impact score, namely $IS(AMB_{i,l})_{overall}$, was calculated for each level (*l*) of the *i*-th ABM by taking into account all the collected information (i.e. appropriateness, magnitude, severity and certainty ratings), in order to weigh it and the measured welfare consequences. All the scores given by the experts were standardized between 0 and 1 (EFSA, 2014a, Section 3.1.4, p.42): ordinal values of 0.33, 0.67 and 1were assigned to the 3 classes of appropriateness; while values of 0.20, 0.40, 0.60, 0.80 and 1 were assigned to the 5 classes of magnitude. The 4 classes of severity were standardized between 0 and 1 using the fol- lowing values: 0, 0.33, 0.67 and 1. The degree of certainty was trans- formed in uncertainty as previously described. First of all, an impactscore for each level *l* of ABM *i* was calculated separately for each expert*k*, as follows: *and foot health*" ($IS(H_{20,3})$ = 2.66; 66.5% of the maximum possiblescore) (Table 2). In agreement with this result, EFSA risk assessment for dairy cow welfare in relation to leg and locomotion problems revealedthat the most important management hazards for cows farmed in cu-bicle houses or straw yards were those related to inadequate care andmonitoring of foot health and hygiene (EFSA, 2009d, Section 9.3, p.18).The presence of one or more animals without access to drinking water was rated with the second most high category impact scores for "*metabolic needs*" ($IS(H_{9,2})$ = 1.87), and "*behavior*" ($IS(H_{9,5})$ = 1.77)(Table 2). Hazards related to insufficient access to drinking water were also found by EFSA (2009a, Fig. 2 and 18, pp. 25 and 47 – therein –, EFSA, 2009b, Fig. 2 and 18, pp. 27 and 53 – therein –) to be correlated with adverse effects impairing dairy cow behavior and metabolism, but they were judged to have quite low magnitude scores. The results ob$IS(ABM) = appropriateness(ABM) \times magnitude$ tained in the present HC could be justified by the fact that Italian $(ABM\ i,k,l\ i,k,l)^{)\times} severity\ i,k,l\ (ABM_{i,k,l})/uncertainty(ABM_{i,k,l})$ (3) weather is characterized by long periods with very high temperatures,as a consequence animal access to water becomes very important for For each ABM, the severity value of level *l* = 1 was set at 0 by default, resulting in $IS(ABM_{i,k,l})_{l=1} = 0$. Finally, the overall impact score $IS(ABM_{i,l})_{overall}$ for levels *l* = 2 and *l* = 3 of the *i*-th ABM were obtained computing the mean of the individual impact scores provided by the 16 experts according to EFSA (2014a, Section 3.1.4, p.42): $IS(ABM_{i,l})_{overall} = mean\,[IS(ABM_{i,k,l})]$. *k*

3. Results and discussion

1. *Hazard characterization*

In the HC exercise, experts' opinion was elicited to rate the negative welfare consequences potentially associated with the exposure of dairy cows to the proposed management and housing hazards. An $IS(H_i)_{overall}$

was determined as in (1) for each of the 52 hazards, allowing to rank them by their relative importance in relation to their potential negative impact on the target population. Results are reported in Figs. 1 and 2 (grey columns). Category impact scores $IS(H_{i,j})$ obtained for each ha- zard (*i*) – welfare category association (*j*), are reported in Tables 2 and 3 (column H, in grey).

Use of electric goads or sharp tools for animal handling (*Type of handling – All*, $IS(H_5)_{overall}$ = 5.25); lack of scheduled foot inspection and trimming and non-use of foot bathing for adult cows (*Foot inspection and foot bathing – Lactating cows, dry cows*, $IS(H_{20})_{overall}$ = 5.05) and presence of one or more animals without access to drinking water (*Water provision – All*, $IS(H_9)_{overall}$ = 4.82) resulted to be the 3 man- agement hazards with the highest potential impact on dairy cow welfare (Fig. 1). The presence of dirty drinkers and dirty water (*Cleanliness of water points*) was considered the factor with the lowest $IS(H_i)_{overall}$ (range 1.04–1.77) regardless of the animal group to which it was re- ferred (*i* = 10–12) (Fig. 1).

The use of electric goads and sharp tools (e.g. pitchfork) for animal handling obtained very high category impact scores for "*behavior*" (IS $(H_{5,5})$ = 2.81) and "*integument alterations*" ($IS(H_{5,4})$ = 1.29) (Table 2), resulting the highest rated hazard within these categories. Use of aversive handling in dairy cows was generally reported in literature to be associated with poor welfare, pain, increase of fear of humans, in- crease of handling problems and increase of injuries to both animals and handlers (Rushen et al., 1999; EFSA, 2009a, Section 13.1, p.19). Thus, avoidance of use of electric goads in cattle was highly re- commended by EFSA (2012b, Appendix A, p.68). In another study carried out by Pajor et al. (2000), the use of electric goads was found to be the most adverse handling method for dairy cattle. Since handling is a necessary and daily procedure in animal husbandry, use of gentle handling methods becomes fundamental in order to prevent several important adverse effects (Rushen et al., 1999).

The lack of scheduled foot inspection and trimming and non-use of foot bathing obtained the highest category impact score for "*locomotion* animal health and welfare. In addition, in the WQ protocol for dairy cattle, thirst was considered more important than hunger for the wel-fare of the animals (Welfare Quality, 2009).

Generally, $IS(H_i)_{overall}$ for the same management hazard, but related to different animal groups, were found to follow the order lactating cows > dry cows > heifers (Fig. 1). Considering housing hazards, the most important ones, due to their potential impact on the welfare of dairy cows kept in loose housing systems, were judged by the 16 experts to be: i) absence or almost absence of bedding material or presence of an inadequate type of bedding material in lactating cow pen (*Type of bedding material – Lactating cows*) ($IS(H_{30})_{overall}$ = 6.61); ii) presence of inadequate or slippery floor in walking areas of lactating cow pen (*Type of floor in walking areas – Lactating cows*) ($IS(H_{33})_{overall}$ = 5.34) and iii) presence of cubicles or deep litter with wrong dimensions or design in lactating cow pen (*Design of lying area – Lactating cows*) ($IS(H_{29})_{overall}$ = 5.22) (Fig. 2).

The absence or the presence of an insufficient quantity of bedding material in lactating cow pen, that force the animals to lie on hard surfaces, or the presence of an inadequate type of bedding material obtained very high impact values in 4 out of the 5 welfare categories: udder health ($IS(H_{30,1})$ = 1.42), locomotion and foot health ($IS(H_{30,3})$ = 1.49), integument alterations ($IS(H_{30,4})$ = 1.90) and behavior (IS ($H_{30,5}$) = 1.48). In particular, this factor was judged by the experts as the most important hazard associated with injuries in lactating cows (Table 3). Several studies report that cows that lie on hard surfaces show high incidence of hock lesions and swollen carpal joints, thus compromising the freedom from pain and injuries and impairing animal welfare (EFSA, 2009e, Section 6.7, pp.105–106; Kielland et al., 2009; von Keyserlingk et al., 2012; Brenninkmeyer et al., 2013). In particular, the amount of bedding material has been demonstrated to be a sig- nificant risk factor for skin lesions on the hock (Kielland et al., 2009).

The presence of inadequate or slippery floor in walking areas of lactating cow pen obtained the highest category impact score for "*lo- comotion and foot health*" ($IS(H_{33,3})$ = 2.12; 53.0% of the maximum possible score) (Table 3). EFSA (2009d, Section 6.1.4, p.16) found that inadequate floor in walking areas (too slippery, too hard, too rough or injuring) had the most important magnitude of the adverse effects as- sociated with leg and locomotion problems for cows farmed in cubicle houses, in agreement with the results of the present study.

Also for housing hazards we found that experts assigned higher overall impact values first of all to factors able to influence lactating cow welfare, followed by the ones referred to dry cows and heifers. In addition, among the selected housing hazards, 4 were related to calvesand 2 of them, absence of bedding material for newborn calves (*Bedding material for newborn calves in single pens*) and presence of not compliant single pens (*Space availability for calves up to 8 weeks*), obtained IS $(Hi)_{overall}$, with i = 43, 44, greater than 3 (Fig. 2).

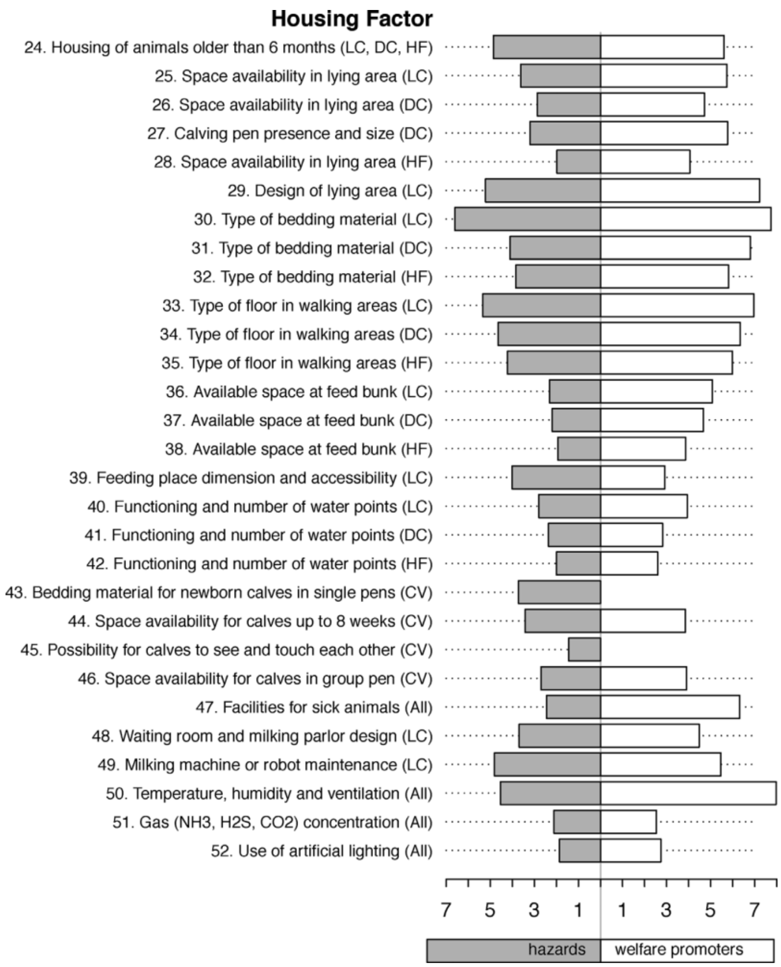

Fig. 2. Overall impact scores of housing hazards ($IS(H_i)_{overall}$)and welfare promoters ($IS(WP_i)_{overall}$).

All = all animal groups; CV = calves; DC = dry cows;HF = heifers; LC = lactating cows.

2. *Welfare promoter characterization*

WPC was carried out to rate the importance of the potential positive welfare consequences associated with the exposure of the target po- pulation to the proposed management and housing welfare promoters. $IS(WP_i)_{overall}$ are reported in Figs. 1 and 2 (white column), while the impact scores referred to the single welfare categories, $IS(WP_{i,j})$, are reported in Tables 2 and 3 (column WP, in white). Considering management welfare promoters, the maintenance of clean bedding material, topped up and renewed regularly, in lactating cow pen (*Bedding material management – Lactating cows*), was found to be the management factor with the most important potential positive impact on dairy cow welfare (*IS*(WP_{16})$_{overall}$ = 6.44). Among the first 3 ranking positions were also classified: presence of experienced and educated stockpersons that take care of animals (*Experience and training of stockpersons – All*) (*IS*(WP_2)$_{overall}$ = 6.17) and implementation of an accurate grouping strategy that meets animal needs in the different life stages (*Animal grouping strategy – Lactating cows, dry cows, heifers*) (*IS* (WP_3)$_{overall}$ = 6.12) (Fig. 1).

An excellent management of the bedding material in lactating cow pen reached the highest score in the category "*integument alterations*" (*IS* ($WP_{16,4}$) = 1.29), and it was found to be associated with quite im- portant positive consequences also in the categories "*udder health*" (*IS* ($WP_{16,1}$) = 2.10), "*locomotion and foot health*" (*IS*($WP_{16,3}$) = 1.37) and "*behavior*" (*IS*($WP_{16,5}$) = 1.13); the recognized benefit intensity of ex- posure was found to positively affect the same welfare categories impaired instead by the hazard level of the same factor (*Bedding material management – Lactating cows*) (Table 2).

Welfare promoters "*experience and training of the stockpersons*" and "*animal grouping strategy*" obtained the two highest category impact score for "*behavior*", *IS*($WP_{2,5}$) = 1.68 and *IS*($WP_{3,5}$) = 1.92, respectively (Table 2). These factors have been reported in literature to bevery important for the welfare of dairy cows (FAWC, 1997; Grant and Albright, 2001), in fact well-trained stockpersons can favor a good human-animal relationship and can promptly recognized early signs of diseases (FAWC, 1997); moreover an accurate grouping strategy can help to fully satisfy animal needs in their different life stages and can reduce social stress during grouping or re-grouping (Grant and Albright, 2001).

Analyzing the *IS*(WP_i)$_{overall}$ given to housing welfare promoters, the most important ones were determined to be: i) adequate temperature, humidity and ventilation, regulated by an automatic cooling system or possibility of summer pasture equipped with shelters (*Temperature, humidity and ventilation – All*) (*IS*(WP_{50})$_{overall}$ = 7.97); ii) presence of optimal bedding material (plentiful, not-abrasive, adequately pre- served, absorbent) in lactating cow pen (*Type of bedding material – Lactating cows*) (*IS*(WP_{30})$_{overall}$ = 7.74) and iii) cubicles or lying area fully occupied by lactating cows thanks to correct dimensions and Category impact scores (*IS*($H_{i,j}$), *IS*($WP_{i,j}$)) of the management factors (*i*) in relation to the 5 identified categories of welfare consequences (*j*). For intensity specification of hazards (*Hi*) and welfare promoters (*WPi*) see Supplementary Table S2.

i	Management Factor (Animal group)	Category of Welfare Consequences (*j*)									
		Udder health (j=1)		Metabolic needs (j=2)		Locomotion and foot health (j=3)		Integument alterations (j=4)		Behavior (j=5)	
		$IS(H_{i,j})$ - $IS(WP_{i,j})$									
		H	WP	H	WP	H	WP	H	WP	H	WP
1	Number of stockpersons (All)	0.83	1.20	0.54	0.97	0.75	1.13	0.40	0.41	0.38	0.65
2	Experience and training of stockpersons (All)	1.14	1.47	0.79	1.21	0.78	1.25	0.20	0.56	1.29	1.68
3	Animal grouping strategy (LC, DC, HF)	0.50	1.18	1.66	1.99	0.39	0.63	0.33	0.40	1.22	1.92
4	Inspections of the animals (All)	0.48	0.63	0.61	0.91	0.96	1.26	0.21	0.58	0.37	0.89
5	Type of handling (All)	0.35	-	0.06	-	0.74	-	1.29	-	2.81	-
6	Milking parlor access and exit (LC)	0.44	-	0.06	-	1.19	-	1.03	-	1.29	-
7	Feeding strategy (LC, DC, HF)	0.82	0.96	2.00	2.41	1.18	1.50	0.03	0.05	0.14	0.46
8	Use of concentrate feeds (daily dose) (LC)	0.83	-	1.65	-	1.29	-	0.02	-	0.09	-
9	Water provision (All)	0.83	0.76	1.87	2.03	0.33	0.16	0.03	0.07	1.77	1.19
10	Cleanliness of water points (LC)	0.51	0.38	0.69	1.14	0.28	0.12	0.06	0.06	0.24	0.73
11	Cleanliness of water points (DC)	0.28	0.31	0.72	1.15	0.10	0.18	0.04	0.06	0.15	0.69
12	Cleanliness of water points (HF)	0.20	0.25	0.51	1.10	0.10	0.03	0.04	0.05	0.20	0.64
13	Cleanliness of floor in walking areas (LC)	0.70	0.71	0.28	0.35	1.91	2.15	0.53	0.89	1.06	1.20
14	Cleanliness of floor in walking areas (DC)	0.47	0.44	0.38	0.30	1.34	1.87	0.45	0.80	0.99	1.03
15	Cleanliness of floor in walking areas (HF)	0.31	0.38	0.25	0.28	1.03	1.75	0.42	0.89	0.94	0.83
16	Bedding material management (LC)	2.03	2.10	0.34	0.56	0.70	1.37	0.73	1.29	0.67	1.13
17	Bedding material management (DC)	1.43	1.47	0.38	0.45	0.59	1.06	0.56	1.28	0.76	1.05
18	Use and management of calving pens (DC)	0.70	1.47	1.10	1.51	0.41	0.50	0.32	0.77	1.34	1.44
19	Bedding material management (HF)	0.87	0.73	0.28	0.30	0.53	0.53	0.57	1.00	0.82	0.98
20	Foot inspection and foot bathing (LC, DC)	0.22	0.36	0.57	0.63	2.66	2.83	0.43	0.35	1.17	1.68
21	Hygiene of the milking parlor or robot (LC)	2.01	1.72	0.23	0.50	0.05	0.28	0.04	0.33	0.04	0.31
22	Milking routine (LC)	2.58	2.52	0.29	0.53	0.03	0.04	0.53	0.56	0.77	1.31
23	Biosecurity measures (All)	1.18	1.69	0.60	0.64	0.46	0.73	0.19	0.61	0.38	0.51

All = all animal groups; DC = dry cows; HF = heifers; LC = lactating cows. design (*Design of lying area – Lactating cows*) ($IS(WP_{29})_{overall}$ = 7.23) (Fig. 2).

Automatic control of temperature, humidity and ventilation or possibility of summer pasture equipped with shelters was determined to be the welfare promoter with the highest positive impact on dairy cow welfare. In particular, it was rated with the highest category impact scores for "*metabolic needs*" ($IS(WP_{50,2})$ = 2.48) and "*behavior*" ($IS(WP_{50,5})$ = 2.40) (≥60.0% of the maximum possible score), but it obtained a high rating also in the category "*udder health*" ($IS(WP_{50,1})$ = 1.54) (Table 3). The control of ambient temperature, humidity and ventilation is in fact considered to have an important role in ensuring a good animal welfare state and in reducing disease prevalence and mortality, especially in Italy, which is characterized by very hot summer months (EFSA, 2009e, Section 6.3.1.1, p.83; Vitali et al., 2009).

Two out of the top 3 housing welfare promoters were referred to dairy cow comfort around resting: type of bedding material was judged very important for enhancing the welfare of lactating cows and design of the lying area obtained high impact values in all the 5 proposed welfare categories (Table 3). Dairy cows have a strong motivation to lie down (EFSA, 2009d, Section 6.6, p.17) and deprivation of resting due to inadequate housing condition was demonstrated to be associated with a state of chronic stress, decrease of rumination time and decrease of blood flow to the udder, thus compromising cow comfort, well-being, production and promoting frustration and abnormal behaviors (Haley et al., 2000).

As already seen for the hazards, $IS(WP_i)_{overall}$ of the same welfare promoter, but related to different animal groups, were higher for lactating cows than for dry cows and heifers, both for management and housing welfare promoters.

3. *Animal-based measure weighing*

The 18 proposed ABMs were firstly judged for their appropriateness by the experts. Using the mean as data aggregation method (EFSA, 2014a, Section 3.1.4, p.42), the 3 most appropriate ABMs for measuring dairy cow welfare outcomes resulted to be: observation of lameness (0.94), annual mortality rate of calves (0.88) and annual mortality rate of adult cows (0.88) (Fig. 3). Observation of lameness was reported to be a high valuable measure for welfare assessment in dairy herds also by Lievaart and Noordhuizen (2011). Moreover, in the same study, the method proposed by Whay et al. (2003a) was found to be one of the two highest scored methods for an effective assessment of lameness cases in dairy cows (Lievaart and Noordhuizen, 2011). This method was the same one chosen in the present study and taught to the experts. Little information is instead available about using mortality as a measure of welfare, however it is considered an indicator characterized by high validity and robustness (Nielsen et al., 2014), which is already routinely collected on-farm in different countries (Nyman et al., 2011; Nielsen et al., 2014). The ABM "*cleanliness of flank, leg, udder*", regardless of the animal group to which it was referred (i.e. heifers, lactating cows or dry cows), obtained the lowest score of appropriateness (range 0.40–0.48) (Fig. 3). In fact, more than 60% of the experts assigned the lowest possible score and nobody gave the maximum score. Lievaart and Noordhuizen (2011) obtained for cow cleanliness a high utility score but a very large range of variability in experts' answers. In the WQ

Table 3

Category impact scores ($IS(H_{i,j})$, $IS(WP_{i,j})$) of the housing factors (i) in relation to the 5 identified categories of welfare consequences (j). For intensity specification of hazards (Hi) andwelfare promoters (WPi) see Supplementary Table S2.

i	Housing Factor (Animal group)	Category of Welfare Consequences (*j*)									
		Udder health (*j*=1)		Metabolic needs (*j*=2)		Locomotion and foot health (*j*=3)		Integument alterations (*j*=4)		Behavior (*j*=5)	
		$IS(H_{i,j})$ - $IS(WP_{i,j})$									
		H	WP	H	WP	H	WP	H	WP	H	WP
24	Housing of animals older than 6 months (LC, DC, HF)	0.20	0.35	0.45	0.94	1.16	1.68	1.05	0.67	2.00	1.98
25	Space availability in lying area (LC)	0.78	1.21	0.40	0.98	0.93	0.98	0.49	0.81	1.02	1.77
26	Space availability in lying area (DC)	0.49	0.83	0.51	0.92	0.60	0.84	0.39	0.69	0.88	1.44
27	Calving pen presence and size (DC)	0.61	1.09	0.87	1.11	0.38	0.74	0.38	0.69	0.97	2.15
28	Space availability in lying area (HF)	0.31	0.59	0.19	0.69	0.45	0.68	0.39	0.68	0.64	1.42
29	Design of lying area (LC)	0.67	1.40	0.51	1.09	1.32	1.35	1.29	1.41	1.42	1.98
30	Type of bedding material (LC)	1.42	1.86	0.33	0.48	1.49	1.52	1.90	1.89	1.48	1.99
31	Type of bedding material (DC)	0.70	1.06	0.23	0.50	0.87	1.47	1.32	1.71	0.98	2.07
32	Type of bedding material (HF)	0.58	0.71	0.06	0.41	1.00	1.29	1.18	1.61	1.02	1.80
33	Type of floor in walking areas (LC)	0.35	0.67	0.40	0.84	2.12	2.14	1.00	1.18	1.48	2.12
34	Type of floor in walking areas (DC)	0.11	0.65	0.60	0.77	1.76	1.95	0.92	1.13	1.25	1.84
35	Type of floor in walking areas (HF)	0.11	0.50	0.29	0.60	1.51	1.95	0.94	1.13	1.36	1.81
36	Available space at feed bunk (LC)	0.17	0.47	0.89	1.62	0.29	0.64	0.16	0.57	0.80	1.77
37	Available space at feed bunk (DC)	0.09	0.43	0.83	1.49	0.29	0.60	0.28	0.55	0.70	1.59
38	Available space at feed bunk (HF)	0.13	0.33	0.63	1.04	0.23	0.47	0.29	0.54	0.66	1.47
39	Feeding place dimension and accessibility (LC)	0.27	0.28	0.89	0.38	0.43	0.21	1.08	0.79	1.35	1.26
40	Functioning and number of water points (LC)	0.31	0.56	0.84	1.76	0.21	0.26	0.27	0.14	1.17	1.22
41	Functioning and number of water points (DC)	0.18	0.33	0.82	1.09	0.14	0.17	0.26	0.11	0.96	1.12
42	Functioning and number of water points (HF)	0.11	0.29	0.60	0.91	0.12	0.14	0.26	0.24	0.91	1.02
43	Bedding material for newborn calves (less than 2 weeks of age) in single pens (CV)	0.00	-	0.39	-	0.66	-	1.23	-	1.43	-
44	Space availability for calves up to 8 weeks (CV)	0.00	0.02	0.33	0.51	0.62	0.42	0.90	0.94	1.58	1.97
45	Possibility for calves to see and touch each other (in single pens) (CV)	0.00	-	0.06	-	0.01	-	0.09	-	1.28	-
46	Space availability for calves in group pen (CV)	0.01	0.03	0.17	0.69	0.52	0.50	0.63	0.81	1.37	1.89
47	Facilities for sick animals (All)	0.47	1.17	0.51	1.47	0.48	1.21	0.26	1.00	0.73	1.48
48	Waiting room and milking parlor design (LC)	0.80	1.28	0.33	0.60	0.64	0.60	0.44	0.25	1.49	1.75
49	Milking machine or robot maintenance (LC)	2.41	2.95	0.41	0.42	0.02	0.10	0.64	0.68	1.33	1.33
50	Temperature, humidity and ventilation (All)	0.62	1.54	1.30	2.48	0.45	1.20	0.09	0.35	2.08	2.40
51	Gas (NH_3, H_2S, CO_2) concentration (All)	0.07	0.35	0.57	0.57	0.15	0.13	0.03	0.14	1.28	1.36
52	Use of artificial lighting (All)	0.12	0.10	0.41	0.82	0.23	0.23	0.24	0.36	0.86	1.24

All = all animal groups; CV = calves; DC = dry cows; HF = heifers; LC = lactating cows.

protocol for dairy cattle, cleanliness was considered less important than measures of behavior for assessing cow comfort around resting (Welfare Quality, 2009). Moreover, in other ruminant species, such as sheep, coat cleanliness was characterized by moderate validity but high fea-sibility (EFSA, 2014a, Table 17, p.61).

Considering magnitude scores, the highest mean values were as- signed to the adverse effects measured by the same ABMs rated as the most appropriate ones: lameness (0.95), annual mortality rate of adult cows (0.80) and annual mortality rate of calves (0.78) (data reported in Supplementary Table S5). Cleanliness of the animals obtained instead the lowest magnitude scores (range 30.0%–41.3% of the maximum possible score). The opinion of the experts confirmed the severity of the adverse effect lameness, which is widely considered to be one of themost serious welfare problem for dairy cows, able to cause long-term pain and consequences on feed intake, milk production and cow re- production (Ito et al., 2010). Also mortality was judged by the expertsas a very important dairy cow welfare issue, as also reported by the scientific literature, which considers dairy cow mortality one of the worst adverse effects (Nielsen et al., 2014), with animals suffering for longer or shorter periods before dying or being euthanized (Thomsen et al., 2006; Fusi et al., 2017).

The $IS(ABM_{i,l})_{overall}$ for levels l = 2 and l = 3 of each ABM, calcu- lated considering all the values given by the experts, are reported in Table 4. Herd-level prevalence of 4–8% of lame adult cows (i = 13, l = 2), 4–10% of dead calves/year ($i = 17$, l = 2) and 15–30% of lactating cows with integument alterations ($i = 10$, l = 2) were con- sidered to have a very important impact on animal welfare (see column "$IS(ABM_{i,l})_{overall}$, $l = 2$" in Table 4). All the ABMs increased their overall impact score moving from level 2 to level 3 (Table 4). In particular the overall impact score of the ABM "*mutilations*", $IS(ABM_{18,l})_{overall}$, in-creased by 178.6% when mutilations are carried out against the national law (e.g. not allowed mutilations, disbudding done in calves older than three weeks of age, mutilations carried out without the su- pervision of a veterinarian) ($IS(ABM_{18,3})$ = 1.56), instead of being done

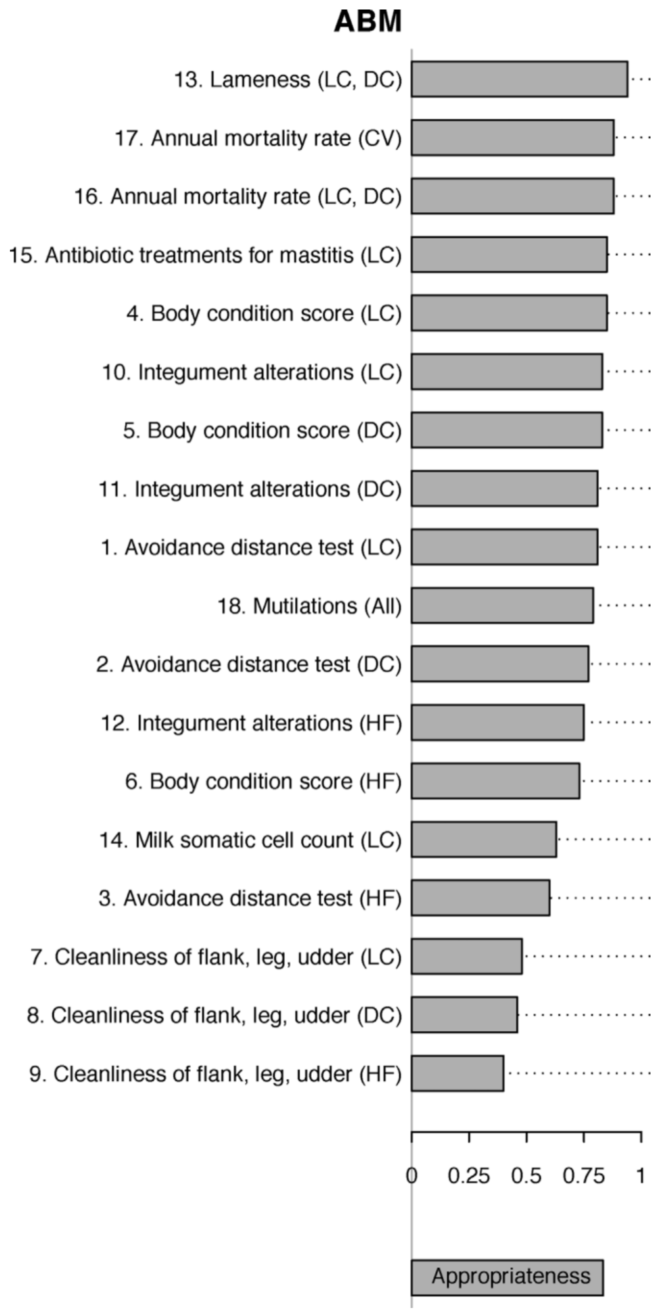

Fig. 3. Mean appropriateness score assigned by the experts to the proposed animal-based measures (ABM). Values given by the experts were standardized between 0 and 1. All = all animal group; CV = calves; DC = dry cows; HF = heifers; LC = lactating cows. according to the law ($IS(ABM_{18,2})$ = 0.56). ABMs that showed the highest overall impact scores at level l = 2 confirmed their importance even at level l = 3; but also the presence of more than 10% of very lean or very fat dry cows ($i = 5, l = 3$) and an annual mortality rate of adult cows greater than 5% ($i = 16, l = 3$) were determined to have a strong impact on the welfare of the herd (see column "$IS(ABM_{i,l})_{overall}$, $l = 3$" in Table 4). The lowest $IS(ABM_{i,l})_{overall}$ were obtained by measures of cleanliness of the animals, independently of the level specifications (Table 4 – $IS(ABM_{i,l})$ with i = 7–9). These results are in agreement with the data reported by Whay et al. (2003b) that carried out an expert opinion elicitation to identify the most

important ABMs for the as- sessment of the welfare state of dairy cattle. In their study, observation of lameness obtained the highest ranking of importance, but also body condition score, observation of limb lesions and examination of mor- tality records obtained high importance scores (> 70% of the maximum Table 4 Overall impact scores ($IS(ABM_{i,l})_{overall}$) of the 18 selected animal-based measures ($ABMi$) at level 2 (l = 2) and 3 (l = 3). For level specifications see Supplementary Table S3.

i ABM (Animal group)	$IS(ABM_{i,l})$, l = 2			$IS(ABM_{i,l})$, l = 3		
	Mean	Min	Max	Mean	Min	Max
1 Avoidance distance test (LC)	0.77	0.00	2.13	1.37	0.09	3.20
2 Avoidance distance test (DC)	0.66	0.00	2.13	1.23	0.09	3.20
3 Avoidance distance test (HF)	0.37	0.00	1.60	0.62	0.09	2.40
4 Body condition score (LC)	0.89	0.00	4.00	1.72	0.09	4.00
5 Body condition score (DC)	1.03	0.00	4.00	1.93	0.09	4.00
6 Body condition score (HF)	0.82	0.00	4.00	1.32	0.09	4.00
7 Cleanliness of flank, leg, udder (LC)	0.20	0.00	1.07	0.39	0.00	1.60
8 Cleanliness of flank, leg, udder (DC)	0.23	0.00	1.07	0.38	0.00	1.60
9 Cleanliness of flank, leg, udder (HF)	0.13	0.00	1.07	0.19	0.00	1.60
10 Integument alterations (LC)	1.08	0.00	4.00	1.63	0.00	4.00
11 Integument alterations (DC)	0.89	0.00	4.00	1.39	0.00	4.00
12 Integument alterations (HF)	0.71	0.00	2.40	1.19	0.00	4.00
13 Lameness (LC, DC)	1.66	0.36	4.00	2.69	0.67	4.00
14 Milk somatic cell count (LC)	0.47	0.00	1.42	0.82	0.00	2.13
15 Antibiotic treatments for mastitis (LC)	1.04	0.13	2.67	1.79	0.27	3.20
16 Annual mortality rate (LC, DC)	1.02	0.09	2.67	1.91	0.18	4.00
17 Annual mortality rate (CV)	1.22	0.27	2.67	2.03	0.53	4.00
18 Mutilations (All)	0.56	0.00	2.67	1.56	0.09	4.00

All = all animal groups; CV = calves; DC = dry cows; HF = heifers; LC = lactating cows.

possible score) (Whay et al., 2003b). In the same study, observation of cow cleanliness was instead placed at the bottom half of the ranking of importance, despite being judged as a suitable and feasible indicator of welfare issues (Whay et al., 2003b).

4. Conclusions

In the present study, the welfare of dairy cows farmed in loose housing systems was found to be greatly affected by adverse effects that result in locomotion and foot problems, death, injuries and serious al- terations of the body condition. As a consequence, ABMs able to properly measure these kinds of effects were judged by the experts to be very important for dairy cow welfare assessment. In addition, experts recognized a great impact to management and housing factors that could be associated with these kinds of outcomes. In particular hazards and welfare promoters linked to comfort around resting, such as man- agement and type of bedding material and design of the lying area, or factors that deal with type of floor in walking areas or with foot in- spection, trimming and foot bathing resulted to be the most important ones. Moreover, several factors were judged to have considerable im- pact on dairy cow behavior, whose assessment and measurement, however, are still an open research topic.

The results obtained in this study have been used as starting point for building the first Italian welfare assessment protocol for screeningthe level of welfare of dairy cows throughout the national territory andto compared Italian herds. The protocol aims at supporting official controls, collecting data, promoting the implementation of on-farm assessment of dairy cow welfare and providing consumers with in- formation, in order to improve the welfare of dairy cows farmed in Italy.

Acknowledgments

This work was supported by the Italian Ministry of Health grant RUMINANTWELFARE. The authors thank the veterinarians that took part to the focus group and to the expert opinion elicitation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in theonline version,.

References

Andreasen, S.N., Sandoe, P., Forkman, B., 2014. Can animal-based welfare assessment be simplified? A comparison of welfare quality protocol for dairy cattle and the simpler and less time consuming protocol developed by the Danish Cattle Federation. Anim. Welf. 23, 81–94.

Aspinall, W., 2010. A route to more tractable expert advice. Nature 463, 294–295.

Boissy, A., Manteuffel, G., Jensen, M.B., Moe, R.O., Spruijt, B., Keeling, L.J., Winckler, C.,Forkman, B., Dimitrov, I., Langbein, J., Bakken, M., Veissier, I., Aubert, A., 2007. Assessment of positive emotions in animals to improve their welfare. Physiol. Behav.

92, 375–397.

Brenninkmeyer, C., Dippel, S., Brinkmann, J., March, S., Winckler, C., Knierima, U., 2013.Hock lesion epidemiology in cubicle housed dairy cows across two breeds, farming systems and countries. Prev. Vet. Med. 109, 236–245.

Broom, D.M., 2017. Animal Welfare in the European Union. European Parliament,

Directorate General for Internal Policies. PE 583.114. http://www.europarl.europa.eu/RegData/etudes/STUD/2017/583114/IPOL_STU(2017)583114_EN.pdf. (Accessed 2 July 2017).

D. Lgs. n. 126/2011. Decreto Legislativo n. 126 del 7 luglio 2011, Attuazione della dir-ettiva 2008/119/CE che stabilisce le norme minime per la protezione dei vitelli. Gazzetta Ufficiale 180 4/8/2011.

D. Lgs. n. 146/2001. Decreto Legislativo n. 146 del 26 marzo 2001, Attuazione della direttiva 98/58/CE relative alla protezione degli animali negli allevamenti. GazzettaUfficiale 95 24/4/2001.

de Vries, M., Engel, B., den Uijl, I., van Schaik, G., Dijkstra, T., de Boer, I.J.M., Bokkers,E.A.M., 2013. Assessment time of the Welfare Quality® protocol for dairy cattle.

Anim. Welf. 22, 85–93.

EFSA, 2009a. Scientific opinion of the panel on Animal Health and Welfare on a request from the Commission on the risk assessment of the impact of housing, nutrition and feeding, management and genetic selection on behaviour, fear and pain problems in

dairy cows. EFSA J. 1139, 1–68.

EFSA, 2009b. Scientific opinion of the panel on Animal Health and Welfare on a request from the Commission on the risk assessment of the impact of housing, nutrition and feeding, management and genetic selection on metabolic and reproductive problems

in dairy cows. EFSA J. 1140, 1–75.

EFSA, 2009c. Scientific opinion of the panel on Animal Health and Welfare on a request from the Commission on the risk assessment of the impact of housing, nutrition and

feeding, management and genetic selection on udder problems in dairy cows. EFSA J. 1141, 1–60.

EFSA, 2009d. Scientific opinion of the panel on Animal Health and Welfare on a request from the Commission on the risk assessment of the impact of housing, nutrition and feeding, management and genetic selection on leg and locomotion problems in dairy

cows. EFSA J. 1142, 1–57.

EFSA, 2009e. Scientific report of EFSA prepared by the Animal Health and Animal Welfare Unit on the effects of farming systems on dairy cow welfare and disease.Annex EFSA J. 1143, 1–38.

EFSA, 2012a. Scientific opinion of the panel on Animal Health and Welfare: guidance on risk assessment for animal welfare. EFSA J. 10 (1), 2513.

EFSA, 2012b. Scientific opinion of the panel on Animal Health and Welfare on the use of animal-based measures to assess welfare of dairy cows. EFSA J. 10 (1), 2554.

EFSA, 2012c. Statement of the panel on Animal Health and Welfare on the use of animal- based measures to assess the welfare of animals. EFSA J. 10 (6), 2767.

EFSA, 2014a. Scientific opinion of the panel on Animal Health and Welfare on the welfare risks related to the farming of sheep for wool, meat and milk production. EFSA J. 12 (12), 3933.

EFSA, 2014b. Guidance on expert knowledge elicitation in food and feed safety risk as- sessment. EFSA J. 12 (6), 3734.

European Commission, 2016. Special Eurobarometer 442 Attitudes of Europeans TowardsAnimal Welfare. https://doi.org/10.2875/884639. (Accessed 02.05.17).

European Union Council, 1998. Council Directive 98/58/EC of 20 July 1998 concerning the protection of animals kept for farming purposes. Off. J. Eur. Communities L 221, 23–27.

FAWC (Farm Animal Welfare Council), 1997. Report on the Welfare of Dairy Cattle. . http://webarchive.nationalarchives.gov.uk/20110909181140/http://www.fawc. org.uk/reports.htm. (Accessed 12 September 2017).

Fusi, F., Angelucci, A., Lorenzi, V., Bolzoni, L., Bertocchi, L., 2017. Assessing circum- stances and causes of dairy cow death in Italian dairy farms through a veterinary practice survey (2013–2014). Prev. Vet. Med. 137, 105–108.

Grandin, T., 2014. Animal welfare and society concerns finding the missing link. Meat Sci. 98, 461–469.

Grant, R.J., Albright, J.L., 2001. Effect of animal grouping on feeding behavior and intakeof dairy cattle. J. Dairy Sci. 84 (E. Suppl), E156–E163.

Haley, D.B., Rushen, J., de Passillé, A.M., 2000. Behavioural indicators of cow comfort: activity and resting behaviour of dairy cows in two types of housing. Can. J. Anim.Sci. 80 (2), 257–263.

Heath, C.A.E., Lin, Y., Mullan, S., Browne, W.J., Main, D.C.J., 2014. Implementing Welfare Quality® in UK assurance schemes: evaluating the challenges. Anim. Welf. 23(1), 95–107.

Hultgren, J., Algers, B., Atkinson, S., Ellingsen, K., Eriksson, S., Hreinsson, K., Nordensten, L., Valtari, H., Mejdell, C.M., 2016. Risk assessment of sheep welfare at small-scale slaughter in Nordic countries, comparing with large-scale slaughter. Acta Vet. Scand.58, 34.

ISTAT, 2017. Consistenza del bestiame bovino e bufalino al 1° dicembre 2016. http://agri.istat.it/jsp/dawinci.jsp?q=plB010000010000011001&an=2016&ig=1&ct= 201&id=8A|9A. (Accessed 26 May 2017).

Ito, K., von Keyserlingk, M.A.G., LeBlanc, S.J., Weary, D.M., 2010. Lying behavior as anindicator of lameness in dairy cows. J. Dairy Sci. 93, 3553–3560.

Kielland, C., Ruud, L.E., Zanella, A.J., Østerås, O., 2009. Prevalence and risk factors forskin lesions on legs of dairy cattle housed in freestalls in Norway. J. Dairy Sci. 92, 5487–5496.

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

Lees, N.M., Lievaart, J.J., 2013. Expert opinion on ranking risk factors for subclinicalmastitis using a modified Delphi technique. N. Z. Vet. J. 61 (3), 168–173.

Lievaart, J.J., Noordhuizen, J.P.T.M., 2011. Ranking experts' preferences regarding measures and methods of assessment of welfare in dairy herds using adaptive con-joint analysis. J. Dairy Sci. 94, 3420–3427.

Lundmark, F., Röcklinsberg, H., Wahlberg, B., Berg, C., 2016. Content and structure of Swedish animal welfare legislation and private standards for dairy cattle. Acta Agric.Scand. A Anim. Sci. 66 (1), 35–42.

More, S.J., Hanlon, A., Marchewka, J., Boyle, L., 2017. Private animal health and welfarestandards in quality assurance programmes: a review and proposed framework for critical evaluation. Vet. Rec. 180, 612.

Nielsen, B.H., Angelucci, A., Scalvenzi, A., Forkman, B., Fusi, F., Tuyttens, F., Houe, H.,Blokhuis, H., Sørensen, J.T., Rothmann, J., Matthews, L., Mounier, L., Bertocchi, L.,Richard, M., Donati, M., Nielsen, P.P., Salini, R., de Graaf, S., Hild, S., Messori, S., Nielsen, S.S., Lorenzi, V., Boivin, X., Thomsen, P.T., 2014. External Scientific Reporton the Use of Animal Based Measures for the Assessment of Dairy Cow Welfare. EFSAsupporting publication EN-659.

Nyman, A., Lindberg, A., Sandgren, C.H., 2011. Can pre-collected register data be used toidentify dairy herds with good cattle welfare? Acta Vet. Scand. 53 (1), S8.

Pajor, E.A., Rushen, J., de Passille, A.M.B., 2000. Aversion learning techniques to eval-uate dairy cattle handling practices. Appl. Anim. Behav. Sci. 69, 89–102.

Ribó, O., Serratosa, J., 2009. History and procedural aspects of the animal welfare risk

assessment at EFSA. In: In: Smulders, F.J.M., Algers, B. (Eds.), Welfare of ProductionAnimals: Assessment and Management of Risks–Food Safety Assurance and Veterinary Public Health, vol. 5. Wageningen Academic Publishers, the Netherlands,pp. 305–335.

Rushen, J., Taylor, A.A., de Passille, A.M., 1999. Domestic animals' fear of humans and its

effect on their welfare. Appl. Anim. Behav. Sci. 65, 285–303.

Smulders, F.J.M., 2009. A practicable approach to assessing risks for animal welfare –

methodological considerations. In: In: Smulders, F.J.M., Algers, B. (Eds.), Welfare ofProduction Animals: Assessment and Management of Risks–Food Safety Assurance and Veterinary Public Health, vol. 5. Wageningen Academic Publishers, the Netherlands, pp. 239–274.

Stewart, D.W., Shamdasani, P.N., 1990. Focus group – theory and practice. In: In:

Bickman, L., Rog, D.J. (Eds.), Applied Social Research Methods Series, vol. 20 SagePublications, Newbury Park, California p. 10.

Thomsen, P.T., Kjeldsen, A.M., Sorensen, J.T., Houe, H., Ersboll, A.K., 2006. Herd-level risk factors for the mortality of cows in Danish dairy herds. Vet. Rec. 158, 622–626.

Vitali, A., Segnalini, M., Bertocchi, L., Bernabucci, U., Nardone, A., Lacetera, N., 2009.

Seasonal pattern of mortality and relationships between mortality and temperature-humidity index in dairy cows. J. Dairy Sci. 92, 3781–3790.

von Keyserlingk, M.A.G., Barrientos, A., Ito, K., Galo, E., Weary, D.M., 2012.

Benchmarking cow comfort on North American freestall dairies: lameness, leg in- juries, lying time, facility design, and management for high-producing Holstein dairycows. J. Dairy Sci. 95, 7399–7408.

Walker, M., Díez-León, M., Mason, G., 2014. Animal welfare science: recent publication

trends and future research priorities. Int. J. Comp. Psychol. 27 (1), 80–100.

Welfare Quality, 2009. Welfare Quality® Assessment Protocol for Cattle. Welfare Quality® Consortium, Lelystad, The Netherlands. http://www.welfarequalitynetwork.net/network/45627/9/0/40. (Accessed 10 May 2017).

Whay, H.R., Main, D.C.J., Green, L.E., Webster, A.J.F., 2003a. Assessment of the welfare of dairy cattle using animal-based measurements: direct observations and in- vestigation of farm records. Vet. Rec. 153, 197–202.

Whay, H.R., Main, D.C.J., Green, L.E., Webster, A.J.F., 2003b. Animal-based measures for

the assessment of welfare state of dairy cattle, pigs and laying hens: consensus ofexpert opinion. Anim. Welf. 12, 205–217.

Yeates, J.W., Main, D.C.J., 2008. Assessment of positive welfare: a review. Vet. J. 175,293–300.

Yousuf, M.I., 2007. Using experts' opinions through Delphi technique. Pract. Assess. Res.