Welfare assessment of horses: the AWIN approach

Emanuela Dalla Costa[1*], Francesca Dai[1], Dirk Lebelt[2], Philipp Scholz[2], Sara Barbieri[1], Elisabetta Canali[1], Adroaldo Jose Zanella[3], Michela Minero[1]

[1] Università degli Studi di Milano, Dipartimento di Scienze Veterinarie e Sanità Pubblica, Milan, Italy

[2] Pferdeklinik Havelland / Havelland Equine Hospital, Beetzsee-Brielow, Germany

[3] Universidade de São Paulo, Departamento de Medicina Veterinária Preventiva e Saúde Animal, Pirassununga, Brazil

*Corresponding author: Emanuela Dalla Costa

Università degli Studi di Milano, Dipartimento di Scienze Veterinarie e Sanità Pubblica, via Celoria 10, 20133 Milano, Italy

Tel: +39 0250318033

Fax: +39 0250318030

E-mail: emanuela.dallacosta@unimi.it

Abstract

The EU-funded Animal Welfare Indicators (AWIN) research project (2011–2015) aimed to improve animal welfare through the development of practical on-farm animal welfare assessment protocols. The present study describes the application of the AWIN approach to the development of a welfare assessment protocol for horses (Equus caballus). Its development required the following steps: (i) selection of potential welfare indicators; (ii) bridging gaps in knowledge; (iii) consulting stakeholders; and (iv) testing a prototype protocol on-farm. Compared to existing welfare assessment protocols for other species, the AWIN welfare assessment protocol for horses introduces a number of innovative aspects, such as implementation of a two-level strategy focused on improving on-farm feasibility and the use of electronic tools to achieve standardised data collection and so promote rapid outcomes. Further refinement to the AWIN welfare assessment protocol for horses is needed in order to firstly gather data from a larger reference population and, secondly, enhance the welfare assessment protocol with reference to different horse housing and husbandry conditions.

Introduction

Among domesticated animals, horses (Equus caballus) are one of the most versatile species: they are used in several activities, ranging, for instance, from agriculture to animal-assisted therapy. It follows that assessment of horse welfare is difficult, in particular when based on resource or management indicators. Their housing and management conditions are so heterogeneous throughout Europe that collecting harmonised data that could be used to make a consistent evaluation of their welfare is complicated. For example, animals may be stabled individually or kept in groups, in areas with or without access to paddock/pasture, and in a facility where either a sole person or several horse owners can be present.

The objective of the Animal Welfare Indicators (AWIN) research project, funded by the European Commission in the Seventh Framework Programme, was to improve animal conditions through the development of practical assessment protocols concerning on-farm welfare of several animal species, including horses. The AWIN approach was based on the method defined in the Welfare Quality® research project (Botreau et al 2007; Blokhuis et al 2010; Rushen et al 2011). The Welfare Quality® project presented four animal welfare principles (Good feeding, Good housing, Good health and Appropriate behaviour) and within these principles highlighted twelve distinct but complementary animal welfare criteria (Blokhuis et al 2010). Using the four animal welfare principles, AWIN researchers aimed to develop a harmonised and scientific welfare assessment protocol for horses based on valid, reliable and feasible animal-based indicators.

An animal-based welfare assessment protocol serves as a toolbox from which it is possible to select the range of measures necessary to address the specific objectives of the evaluation for that particular species and category of animal at that time (EFSA Panel on Animal Health and Welfare 2012). Animal-based indicators were chosen because they relate directly to the animal itself rather than to the environment in which the animal/individual is kept (EFSA Panel on Animal Health and Welfare 2012). The indicators can be collected in different housing conditions and used to infer how the animal is affected by external factors.

The present study describes the AWIN approach to the development and on-farm use of a welfare assessment protocol for horses.

The development of the AWIN protocol

The development of the AWIN welfare assessment protocol for horses required the following: i) selection of potential welfare indicators; ii) bridging knowledge gaps; iii) stakeholder consultation; and iv) testing of the model protocol on-farm.

Selection of potential welfare indicators

A group of thirteen academic scientists, renowned internationally for their expertise in equine welfare and authors of peer-reviewed publications on relevant topics, were designated as partners and collaborators in the AWIN project. They systematically reviewed available relevant scientific literature to select promising animal-based indicators for use on-farm in horses (Dalla Costa et al 2014b). Scientists evaluated validity, reliability and on-farm feasibility of each indicator (Scott et al 2001) according to the definitions that were provided them at the beginning of the review process. Indicators were then classified according to the four principles and the twelve criteria developed by the Welfare Quality® project (Blokhuis et al 2010). Following face-toface consultation, scientists agreed on the selection of at least one promising indicator for each welfare criterion to be included in the prototype protocol. The scientists highlighted that certain indicators had been well investigated and could be considered ready for use on-farm while others showed gaps in scientific knowledge (Table 1).

Bridging gaps in knowledge

The next stage was to develop a research action-plan to address lack of knowledge regarding validity, repeatability and feasibility of single potential/promising indicators. In cases where no Gold Standard measures (reference measures widely recognised as being the best available [Versi

4

1992]) existed to address a specific welfare criterion, specific validation studies were carried out (Dalla Costa et al 2014a, 2015; Dai et al 2015). For example, as no Gold Standard indicator has been identified to assess pain in horses, a new measure, the Horse Grimace Scale, was developed and applied to acute, post-surgical castration pain as a standardised model of pain.

Consultation of stakeholders

Following the literature review and research studies, a prototype protocol was developed and subjected to stakeholder consultation with the aim of gathering the views of groups of individuals who share a close interest in the equine sector. Stakeholder input was proactively sought and significant effort was expended to explain the development process of the welfare assessment protocol. The stakeholders' involvement was intended not only to increase acceptability of its outcomes, through stimulation of a multidisciplinary dialogue, but also to identify potential barriers to the practical application of the protocols, and possible related solutions. To this end, a survey was made available in five different languages through the AWIN project and other institutions/organisations (eg Food and Agriculture Organization of the United Nations and Federazione Italiana Sport Equestri) websites for 15 months (December 2012–March 2014). Moreover, stakeholders were engaged in participatory activities designed for community facilitators in collective actions (van Dijk et al 2011) during their face-to-face consultation. For instance, they were asked to compile notes describing signs on a horse's body that indicate good welfare, or fill in matrices making direct comparisons between welfare criteria.

Two hundred and seventy nine people (Owner = 45%; Veterinarian = 19%; Trainer = 13%; Veterinary technician/Nurse = 12%; Farmer = 11%) from different European (67%) and non-European (eg Asia, Australia, North and South America, Europe) countries (33%) participated in the online survey and a further 58 people took part in face-toface interviews. Both the online survey and the consultation involved different stakeholders, such as official veterinarians, NGOs, horse owners, grooms and riders. All stakeholders agreed upon who should be in charge of the animal

welfare assessment as well as its objectives and time needed to perform the assessment. Veterinarians should evaluate equine welfare; whilst owners should be trained to properly assess their own animals' welfare. Participants felt that the average maximum time per horse assessment should range between 5 and 10 min, although a small number of participants (n = 24) reported 30 min. Stakeholders also agreed on the fact that no welfare principle is fully adhered to for horses and, during the face-toface meeting, made direct comparisons between perceived relevance of welfare criteria, in order to prioritise different issues. The Bradley-Terry method (Agresti 2013) was used to model the probability that a given criterion 1 is perceived as more relevant than a criterion 2 (Figure 1).

Stakeholders ranked 'absence of prolonged thirst' and 'appropriate nutrition' as the most important criteria to assess equine welfare whilst 'good human-animal relationship', 'comfort around resting' and the ability to 'express other behaviours' were considered relatively less important. Stakeholder consultation promoted an informed debate on sensitive issues regarding the acceptability of the welfare assessment process.

Testing a prototype protocol on-farm

Forty horse facilities (riding school = 37%; training centre = 24%; breeding farm = 15%; hippodrome = 3%; other (eg animal-assisted activity) = 21%), were visited in Italy and Germany between March and July 2014. All participants took part on a voluntary basis. The farms were sampled according to their geographical distribution and number of horses present. Both in Italy and in Germany, regions or states with the highest number of horses were identified and official databases from local animal health authorities were consulted to determine the number of facilities, the number of horses in each facility, and their geographical location. An inclusive criterion considered horses being stabled indoors for at least 12 h per day, as this was reported to be the most prevalent housing system for horses in different European countries (Søndergaard & Winther Christensen 2002; Knubben et al 2008; Hartman et al 2012; Hockenhull & Creighton 2015).

The analysis revealed the presence of a large number of very small horse facilities, which have the potential for poor welfare through a lack of social contact among individuals. Therefore, a stratified random sample of very small (≤ 4 horses), small (5–10 horses), medium (11–30 horses) and large (> 31 horses) horse facilities was adopted. Finally, the prototype protocol was refined according to the results of scientific studies performed in step 2 (bridging gaps in knowledge), the general comments from the stakeholders and the on-farm testing (AWIN 2015).

The AWIN protocol

The objective of the development of the 'AWIN welfare assessment protocol for horses' was to produce a welfare assessment tool box for single-stabled horses over five years old that could be used by trained assessors, and which should be valid for all horses, regardless of their attitude or the resources provided in terms of housing or management condition.

The protocol uses a two-level approach, beginning with screening designed to identify welfare issues ranked as important by stakeholders. It includes easy-to-apply, valid and reliable welfare indicators; requires no animal handling; takes a short time and requires little training for assessors to apply it. This first level focuses on a rapid response and on a consequent shorter time for the assessment while maintaining its accuracy. Indicators that were included in the first level were 'iceberg indicators', ie a subset of animal-based measures that provide an overall assessment of their welfare (Farm Animal Welfare Council 2009). Depending on the outcome of the first-level assessment, a more comprehensive and in-depth second-level assessment may be recommended and implemented.

First-level welfare assessment protocol

The first-level assessment consists of 18 welfare indicators covering all of the principles developed by Welfare Quality® (Blokhuis et al 2010). Even though some indicators may reflect the existence of more than one issue, a positive outcome of each indicator shows that a specific criterion has been

fulfilled. For example, 'poor Body Condition Score' can be related to a variety of factors such as food availability, disease or feeding practices, however, 'optimal Body Condition Score' would appear a clear reflection of prolonged appropriate nutrition.

Most indicators are animal-based; however, for some criteria (eg absence of prolonged thirst) resource- or management-based measures were included because no animal-based indicators met the required characteristics of validity, reliability or feasibility. Resource-based indicators in the first-level welfare assessment protocol are: box dimensions; water availability; bedding characteristics; and possibility of social interaction between horses. The only management-based measurement consists of a questionnaire concerning the possibility for a horse to spend some time outside its box on a daily basis (exercise). (Detailed information on descriptions, assessments and scoring systems for each indicator can be found in the 'AWIN welfare assessment protocol for horses'; AWIN 2015).

Figure 2 represents the workflow of the first-level welfare assessment: indicators are recorded mostly from outside the box in a logical order to maximise feasibility and minimize observer influence on horse reactions. Approximate time needed for assessing a horse is 5 min.

For a reliable assessment of a farm, it is not necessary to evaluate all the animals that are housed. The number of horses to be sampled and assessed can be determined according to the table reported in the protocol (AWIN 2015), which is calculated for an expected variation in data of 0.5, at a confidence level of 0.9 and a precision of the estimate (δ) of 0.1. Random selection of horses from microchip numbers allows many of the possible sources of bias that could affect animal sampling on-farm to be avoided.

Progression from first- to second-level welfare assessment protocol

The second-level assessment is recommended in the following cases: (i) only one horse is housed at the facility; (ii) a non-compliance with the current legislation exists; (iii) the within-farm proportion of animals meeting a given criterion is lower than the proportion of animals observed for the same

criterion in the fifth percentile of the farms from the reference population. To date, the reference population against which to compare farms is based on the farms assessed during the AWIN project (Dalla Costa et al 2016). In the future, the reference population will be updated taking into consideration a wider geographic area, and a larger number of farms and horses.

Second-level welfare assessment protocol

The second-level assessment consists of 25 welfare indicators thoroughly described in the protocol (AWIN 2015). Eighteen of the welfare indicators are assessed following the same procedure adopted in the first level whilst others, such as the bucket test, coughing, lameness, lesions at mouth corners, the fear test, the forced human approach test and Qualitative Behaviour Assessment (Dalla Costa et al 2012) pertain only to the second-level assessment. Approximate time needed for assessing a single horse varies between 11 and 25 min, depending on individual differences in the reaction to the behaviour tests and on the necessity of a bucket test to evaluate thirst. In the second level all the animals are to be assessed and removed from their stables in order to evaluate lameness, integument alterations, lesions at mouth corners, swollen joints and signs of hoof neglect. Figure 3 reports the flow of the second-level welfare assessment.

Carrying out the second-level welfare assessment is more challenging than performing the first in terms of time and handling of animals; nonetheless, disturbance to the animals and changes in their daily routine are kept to a minimum. The second level is performed when an in-depth investigation on the welfare condition of horses is essential.

Data collection and output

To improve efficiency and reliability of on-farm data collection, reduce transcription mistakes and perform automatic data storage, AWIN researchers, in collaboration with DAIA Intelligent Solutions SL (Ordizia, Spain), developed the AWINHorse app, which is freely available for tablets and smartphones. The AWINHorse app enables the user to collect first-level welfare assessment

data and store information on their own device, delivering an immediate visual output regarding the welfare status of the animals being assessed. Minimal training is needed to use the app, however, no individual or organisation can be considered capable of applying the protocol in a robust and reliable way without appropriate training on how to assess and score the indicators. The app automatically provides an output where data are displayed in bar charts (Figure 4) and the position of the assessed farm is highlighted in comparison with the median value of the reference population. Welfare indicators are aggregated at criterion level (Table 2) and the graph shows the proportions of horses, within the assessed farm, for which the criterion is satisfied. Not only does the app increase efficiency and transparency of the assessment process, but it also stimulates a dialogue with horse owners about the results of the assessment and the actions needed to improve the welfare of their animals. For each farm, data are automatically collated by the app (each horse corresponds to a row of a CSV file) and used to create the output. The application also allows the user to send data to a central server and, should there be a requirement for further analysis, data collected can be downloaded into an Excel file.

In case of a farm housing less than ten horses, it is preferable to consider how any single animal complies with each criterion rather than calculate the proportions of animals. Data can also be collected on the recording sheet enclosed in the protocol (AWIN 2015).

Data gathered using the second-level welfare assessment protocol can be interpreted by calculating the proportion of animals showing different scores for each welfare indicator. As it stands, no application or software is available to automatically calculate an output from the second-level input data.

Adaptation for group-housed horses

In principle, welfare indicators included in the 'AWIN welfare assessment protocol for horses' could be applied to horses stabled in groups. As some indicators might show limitations in the feasibility of their application, the protocol was pilot-tested in ten German facilities where horses

were kept in groups. Eventually, this pilot test resulted in some suggestions for adaptation of the protocol which are reported in the Annex C to AWIN (2015). Shelter dimensions, signs of thermal stress, agonistic behaviour, human-animal relationship tests and Qualitative Behaviour Assessment were adapted and accepted even in anticipation of further improvements and refinements supported by scientific studies.

Animal Welfare Implication

The application of a scientifically sound on-farm welfare assessment protocol accepted by stakeholders is an important milestone on the road leading to the improvement of the quality of equine life.

Conclusion

A major objective in equine research is the development of a scientifically sound process for the assessment of animal welfare that is valid, reliable and able to be implemented successfully at farm level. This implies a combination of fundamental scientific requirements together with applied aspects.

This paper presents the approach adopted in AWIN which enabled scientists to deliver a comprehensive, easy-to-use welfare assessment protocol for horses that includes animal-based indicators derived from scientific literature and/or developed within AWIN and published in peer-reviewed scientific journals.

Though similar to the Welfare Quality® project in terms of its methodology, this AWIN welfare assessment protocol for horses shows unique features, such as a two-level strategy; a dedicated IT application — the AWIN Horse app that generates an immediate output of the assessment — and a systematic collection of standardized data. Moreover, the strategic and participatory approach to stakeholders played an important role in creating constructive relationships and maintaining them

over time. An example of which was the National Competent Veterinary Authorities frequent requirement to apply the 'AWIN welfare assessment protocol for horses' in their everyday work. There are still a number of important challenges to be addressed. To make the reference population more geographically representative, the data collection approach must become larger and be more widely distributed. In anticipation of further scientific research, the welfare assessment protocol includes some suggestions for its adaptation to specific management situations and husbandry types, eg group-stabled horses. Future research is needed to improve the understanding of specificity of single, animal-based indicators, clarifying whether a single indicator can detect changes in the animals' responses, which may be relevant for their welfare status, and whether it is related to a single welfare consequence or responds to several different consequences (EFSA Panel on Animal Health and Welfare 2012). Furthermore, two major issues remain to be investigated to make the welfare assessment more and more effective: how environmental changes can influence a single measure, in order to define the sensitivity of the animal-based indicators to be included, and what is relevant in recent neuroscience research that can be used to interpret horse behaviours in relation to positive or negative affect (an issue of increasing importance in animal welfare research).

References

Agresti A 2013 Categorical Data Analysis, Third Edition. Wiley: New York, USA

AWIN 2015 AWIN welfare assessment protocol for horses. http://dx.doi.org/10.13130/AWIN_horses_2015

Blokhuis HJ, Veissier I, Miele M and Jones B 2010 The Welfare Quality® project and beyond: safeguarding farm animal well-being. Acta Agriculturæ Scandinavica, Section A - Animal Science 60: 129-140

Botreau R, Veissier I, Butterworth A, Bracke MBM and Keeling LJ 2007 Definition of criteria for overall assessment of animal welfare. Animal Welfare 16: 225-228

Dai F, Cogi NH, Ugo E, Heinzl L, Costa ED, Canali E and Minero M 2015 Validation of a fear test in sport horses using infrared thermography. Journal of Veterinary Behavior: Clinical Applications and Research 10(2): 128-136. http://dx.doi.org/10.1016/j.jveb.2014.12.001

Dalla Costa E, Dai F, Lebelt D, Scholz P, Barbieri S, Canali E and Minero M 2016 Initial outcomes of a harmonized approach to collect welfare data in sport and leisure horses. Animal 13: 1-7. http://dx.doi.org/10.1017/S1751731116001452

Dalla Costa E, Dai F, Murray LAM, Guazzetti S, Canali E and Minero M 2015 A study on validity and reliability of on-farm tests to measure human-animal relationship in horses and donkeys. Applied Animal Behaviour Science 163: 110-121. http://dx.doi.org/10.1016/j.applanim.2014.12.007

Dalla Costa E, Dai F, Villa C, Canali E and Minero M 2012 Ongoing studies: Qualitative Behavioural Assessment (QBA) of positive emotional state of stabled horses. First AWIN Annual Conference. 21–24 May 2012, Oscarsborg, Norway

Dalla Costa E, Minero M, Lebelt D, Stucke D, Canali E and Leach M 2014a Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. PloS One 9: e92281. http://dx.doi.org/10.1371/journal.pone.0092281

Dalla Costa E, Murray L, Dai F, Canali E and Minero M 2014b Equine on-farm welfare assessment: a review of animal based indicators. Animal Welfare 23: 323-341. http://dx.doi.org/10.7120/09627286.23.3.323

EFSA Panel on Animal Health and Welfare (AHAW) 2012 Statement on the use of animal-based measures to assess the welfare of animals. EFSA Journal 10: 1-29

Farm Animal Welfare Council 2009 Farm Animal Welfare in Great Britain: Past, Present and Future. FAWC: London, UK

Hartmann E, Søndergaard E and Keeling LJ 2012 Keeping horses in groups: A review. Applied Animal Behaviour Science 136: 77-87. http://dx.doi.org/10.1016/j.applanim.2011.10.004

Hockenhull J and Creighton E 2015 The day-to-day management of UK leisure horses and the prevalence of owner reported stable-related and handling behaviour problems. Animal Welfare 24: 29-36. http://dx.doi.org/10.7120/09627286.24.1.029

Knubben JM, Gygax L and Stauffacher M 2008 Horses in Switzerland: Results of a representative survey of population, housing and use in 2004. Schweizer Archiv für Tierheilkunde 150: 387-397. http://dx.doi.org/10.1024/0036-7281.150.8.387

Rushen J, Butterworth A and Swanson JC 2011 Animal behavior and well-being symposium. Farm animal welfare assurance: science and application. Journal of Animal Science 89: 1219-1222. http://dx.doi.org/10.2527/jas.2010-3589

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 30: 5-10

Søndergaard E and Winther Christensen J 2002 A survey of housing and management conditions for horses in Denmark.http://research.vet.upenn.edu/HavemeyerEquineBehaviorLabHomePage/ReferenceLibraryHavemeyerEquineBehaviorLab/HavemeyerWorkshops/HorseBehaviorandWelfare1316June2002/HorseBehaviorandWelfare2/ASurveyofHousingandManagementConditionsfor/tabid/3133/Default.aspx

Van Dijk L, Pritchard J, Pradhan S and Wells K 2011 Sharing the load: a guide to improving the welfare of working animals through collective action. Practical Action Publishing, Wallingford, UK. http://dx.doi.org/10.3362/9781780440675

Versi E 1992 'Gold standard' is an appropriate term. British Medical Journal 305(6846): 187

Table 1 Promising welfare indicators selected by scientists and divided for welfare criteria. The

column 'actions' reports the plan to cover gaps in scientific knowledge for some of the indicators.

Welfare principles	Welfare criteria	Welfare indicators	Actions
Good feeding	Appropriate nutrition	Body Condition Score	NN
	Absence of prolonged thirst	Resource-based (water availability), bucket test	Evaluate feasibility of bucket test
Good housing	Comfort around resting	Resource-based (bedding, box dimensions)	NN
	Thermal stress	Resources to reduce thermal stress (eg shade, rugs), shivering, increased frequency/depth of respiration, flared nostrils, profuse sweating, apathy	NN
	Ease of movement	Management-based (exercise)	NN
Good health	Absence of physical injuries	Integument alterations, swollen joints, lameness, prolapse	Evaluate feasibility of lameness
	Absence of disease	Hair coat condition, discharges, consistency of manure, abnormal breathing, coughing	NN

	Absence of pain and pain induced by management procedures	Horse Grimace Scale, signs of hoof neglect, lesions at mouth corners, pain-related behaviour, lesions inner lip, tongue and bars	Validation of HGS, evaluate feasibility of pain-related behaviour and lesions at mouth corners and lesions inner lip, tongue and bars
Appropriate behaviour	Expression of social behaviour	Resource-based (social interaction)	NN
	Expression of other behaviours	Stereotypies, fear test	Evaluate validity and feasibility of Fear test
	Good human-animal relationship	Human-animal relationship tests	Evaluate validity and inter-observer reliability of Human-animal behaviour tests
	Positive emotional state	Qualitatitive	Evaluate validity and interobserver reliability of QBA

NN = Not needed. It refers to resource- and management-based indicators or to indicators that have already been validated and tested for reliability and on-farm feasibility.

Table 2 Aggregation of welfare indicators described for each criterion.

WQ® criteria	Aggregation criteria
Appropriate nutrition	Proportion of horses with appropriate Body Condition Score
Absence of prolonged thirst	Proportion of horses with clean fresh water available
Comfort around resting	Proportion of horses with satisfactory box dimensions and sufficient clean bedding
Ease of movement	Proportion of horses with the possibility of enough daily exercise
Absence of physical injuries	Proportion of horses without any physical injuries
Absence of disease	Proportion of horses without any signs of disease
Absence of pain and pain induced by management procedures	Proportion of horses without any signs of pain
Expression of social behaviour	Proportion of horses with the possibility to have social interaction
Expression of other behaviours	Proportion of horses without any stereotypies
Good human-animal relationship	Proportion of horses with positive score to all human-animal relationship tests

Figure 1 The chart reports the probability (cuts) that a given criterion 1 (y-axis) is perceived as more relevant than a criterion 2 (x-axis) by stakeholders.
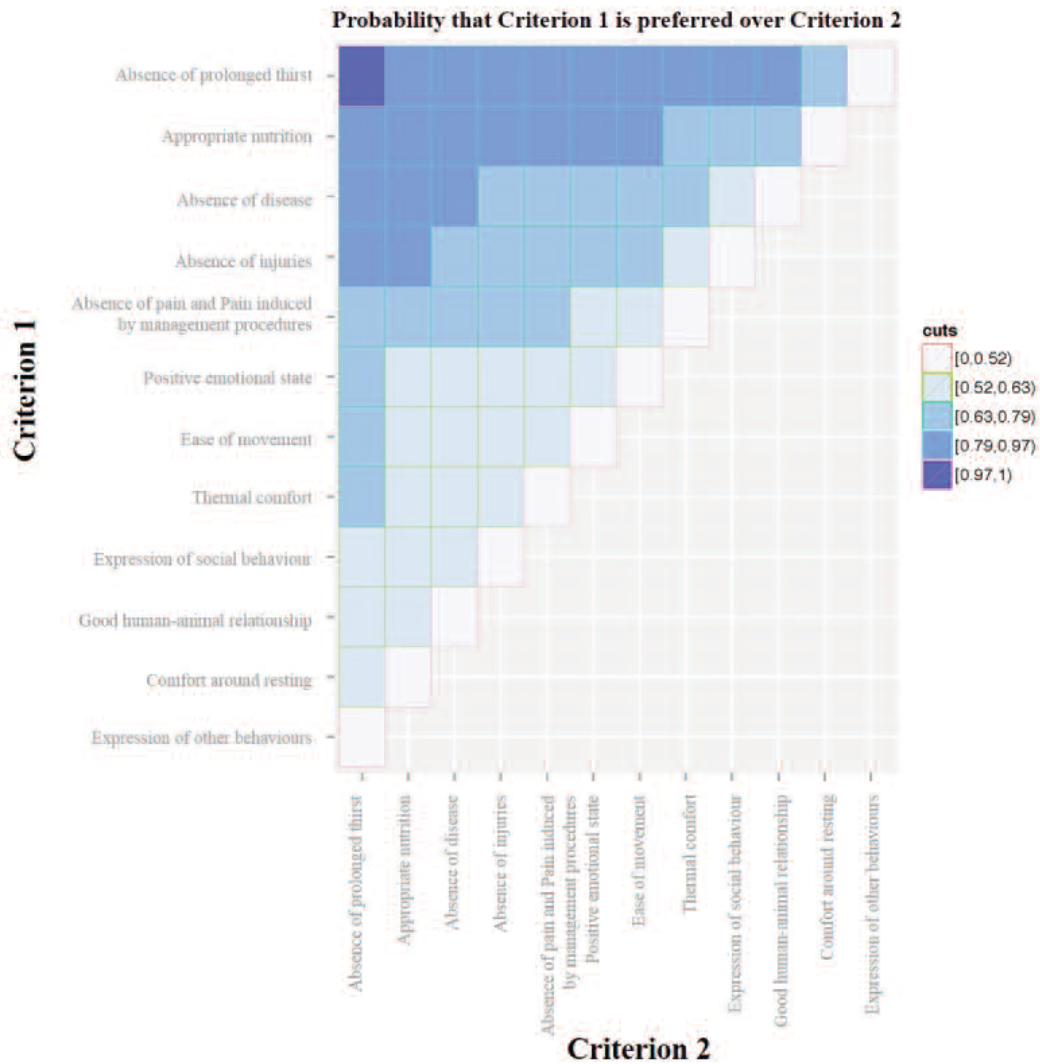
Figure 2 First-level welfare assessment flowchart. The graph represents the specific order in which the first-level welfare indicators should be collected on-farm.
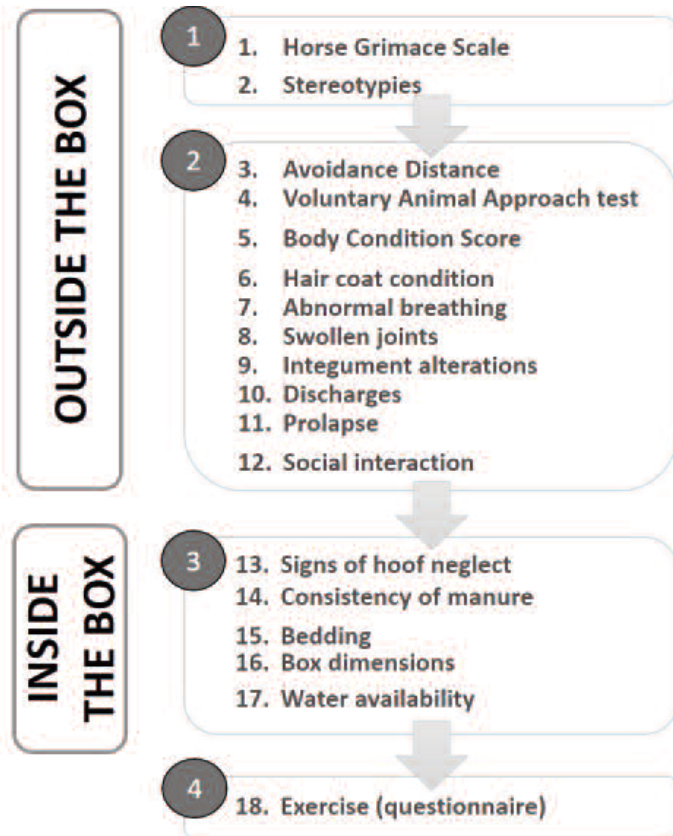
Figure 3 Second-level welfare assessment flow chart. The graph represents the specific order in which the second-level welfare indicators should be collected on-farm.
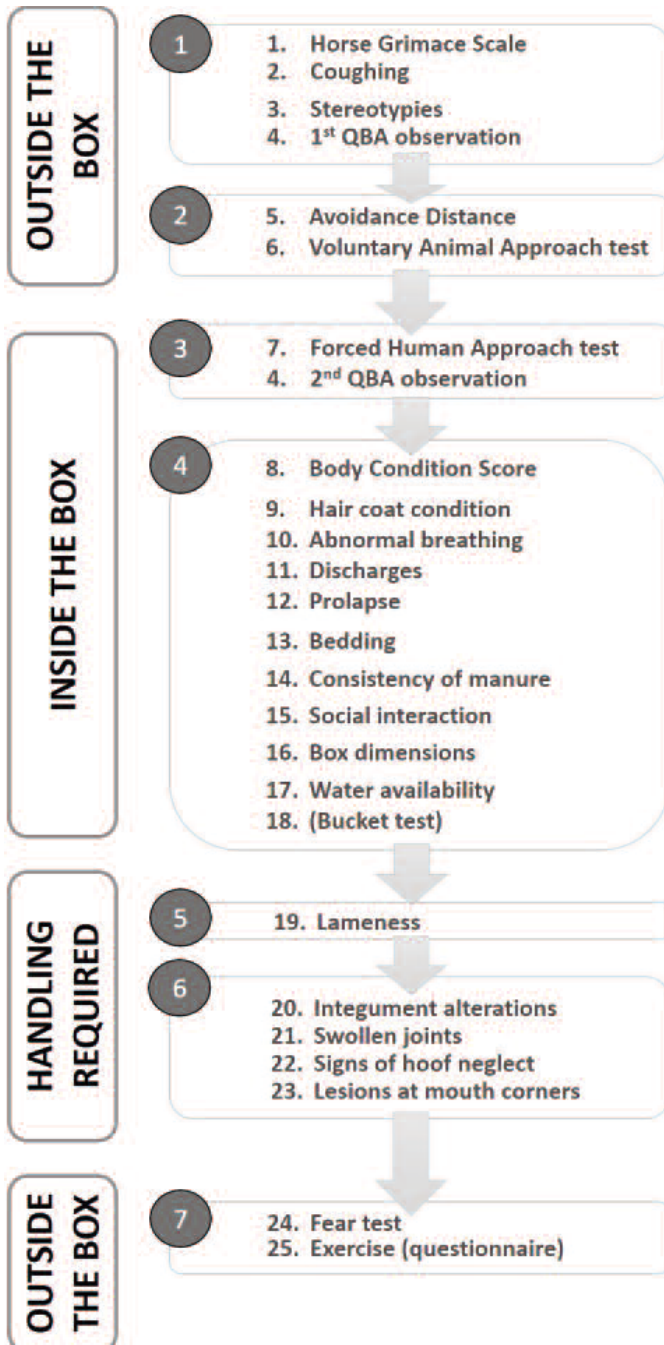
Figure 4 Example of output of the first-level welfare assessment. Information is aggregated at criterion level, resulting in an assessment of how an individual farm complies with each criterion. The criteria are displayed in the output and the position of the assessed farm (the yellow dot) is highlighted in comparison with the median value of the reference population (blue bars).