Evaluation of consistency over time of the use of the Animal Welfare Indicators protocol for horses

I Czycholl*, K Büttner†, P Klingbeil† and J Krieter†

Abstract

Consistency over time is a basic requirement for welfare assessment schemes since consistency must not depend, for example, on the day it is carried out. This study analysed the consistency of the indicators of the Animal Welfare Indicators (AWIN) protocol for horses (Equus caballus) over time. Given the multi-dimensionality of animal welfare, the AWIN protocol includes a variety of indicators evaluating, eg the health status or the behaviour of the animals. Fourteen establishments keeping horses in Germany were visited four times each (day 0, day 3, day 42, day 90). For the evaluation of reliability and agreement between the different visits, ie across time, the reference visit on day 0 was compared to the other visits via calculation of Spearman’s rank correlation (RS), intra-class correlation (ICC), smallest detectable change (SDC) and limits of agreement (LoA). The indicator, Qualitative Behaviour Assessment (QBA) was analysed by Principal Component Analysis (PCA). Most of the indicators demonstrated sufficient consistency over time. Indicators that were inconsistent included parts of the Horse Grimace Scale, outcomes of behavioural tests, the presence of swollen joints as well as the indicators hoof neglect, alopecia on the legs and water cleanliness. The QBA was consistent for the period of 42 days, but not for 90 days. Overall, those indicators with insufficient consistency over time require to be revised or replaced in future welfare assessment schemes.

Keywords: animal-based indicators, animal welfare, AWIN, consistency, horse, reliability

Introduction

Welfare is generally acknowledged to be a multi-dimensional concept made up of good health and biological functioning, natural behaviour and emotional state. Hence, for the accurate assessment of welfare, multiple indicators are required since a single indicator is insufficient to measure all these different aspects (Fraser 2008; Blokhuis et al 2013; Czycholl et al 2015). In general, animal-based indicators, which are also outcome-based indicators, are considered the most useful in terms of assessing the true welfare state during on-farm welfare assessments (Blokhuis et al 2013). In contrast, management- and resource-based indicators constitute a risk assessment of the surroundings and the possible attainment of a good welfare state for the animals. However, the measurement of animal-based indicators poses the greatest challenge as regards feasibility, reliability and validity. For example, behavioural observations to assess animals’ lying behaviour are more time-consuming and require better trained assessors than simply assessing the size of the resting area (Velarde & Geers 2007). Hence, a major issue in animal welfare science continues to be the suitability of animal-based indicators for the purposes of a welfare assessment.

Although generally important for all species in animal welfare assessment, animal-based indicators are probably especially important when it comes to an assessment of equine welfare since the use of this species, not to mention the husbandry conditions, show great variation, eg ranging from leisure horses (Equus caballus) to working equids and food production (Dalla Costa et al 2017). Animal-based indicators allow comparison between different husbandry conditions and make it possible for objective comparison to take place (Dalla Costa et al 2014).

In 2015, a welfare assessment protocol for horses was developed within the framework of the Animal Welfare Indicators (AWIN 2015) project, amongst others. Welfare was defined as a multi-dimensional concept to be measured by a variety of predominantly animal-based indicators. Resource- (eg water provision) or management-based (eg exercise) indicators were only included if no feasible, reliable and valid animal-based indicator was revealed for a certain aspect of animal welfare. However, research concerning the feasibility, reliability and validity of the included indicators remains an ongoing process. Some single indicators were assessed for feasibility, reliability and validity before inclusion in the respective protocols. For example, Dalla Costa et al (2014, 2016b) validated the Horse Grimace Scale as a tool to assess pain by comparing horses in pain to control groups without pain or receiving analgesia. Dai et al (2015) validated the fear tests by...