Comparison of locomotion scoring for dairy cows by experienced and inexperienced raters using live or video observation methods

A Schlageter-Tello*, EAM Bokkers‡, PWG Groot Koerkamp§, T Van Hertem¶, S Viazzi#, CEB Romanini#, I Halachmi¶, C Bahr#, D Berckmans# and K Lokhorst†

1 Wageningen UR Livestock Research, PO Box 338, 6700 AH, Wageningen, The Netherlands
2 Animal Production Systems Group, Wageningen University, PO Box 338, 6700 AH, Wageningen, The Netherlands
3 Farm Technology Group, Wageningen University, PO Box 317, 6700 AH, Wageningen, The Netherlands
* Division Measure, Model & Manage Bioresponses, KU Leuven, PO Box 2456, 3001 Heverlee, Belgium
† Institute of Agricultural Engineering, Agricultural Research Organisation, PO Box 6, Bet-Dagan 50250, Israel
# Contact for correspondence and requests for reprints: andres.schlagetertello@wur.nl

Abstract

Lameness is considered a major problem in dairy production. Lameness is commonly detected with locomotion scores assigned to cows under farm conditions, but raters are often trained and assessed for reliability and agreement by using video recordings. The aim of this study was to evaluate intra- and inter-rater reliability and agreement of experienced and inexperienced raters for locomotion scoring performed live and from video, and to calculate the influence of raters and the method of observation (live or video) on the probability of classifying a cow as lame. Using a five-level locomotion score, cows were scored twice live and twice from video by three experienced and two inexperienced raters for three weeks. Every week different cows were scored. Intra- and inter-rater reliability (expressed as weighted kappa, κw) and agreement (expressed as percentage of agreement, PA) for live/live, live/video and video/video comparisons were determined. A logistic regression was performed to estimate the influence of the rater and method of observation on the probability of classifying a cow as lame in live and video observation. Experienced raters had higher values for intra-rater reliability and agreement for video/video than for live/live and live/video comparison. Inexperienced raters, however, did not differ for intra- and inter-rater reliability and agreement for live/live, live/video and video/video comparisons. The logistic regression indicated that raters were responsible for the main effect and the method of observation (live or from video) had a minor effect on the probability for classifying a cow as lame (locomotion score ≥ 3). In conclusion, under the present experimental conditions, experienced raters performed better than inexperienced raters when locomotion scoring was done from video. Since video observation did not show any important influence in the probability of classifying a cow as lame, video observation seems to be an acceptable method for locomotion scoring and lameness assessment in dairy cows.

Keywords: animal welfare, lameness, live scoring, observers, repeatability, video-scoring

Introduction

Lameness is considered a major problem in dairy production (Bruijnis et al. 2010). Mean prevalence of lameness in dairy herds during the last decade was 33% in Austria and Germany (Dippel et al. 2009a,b); 37% in England and Wales (Barker et al. 2010); and ranged from 21 to 55% in the USA (Cook 2003; Espejo et al. 2006; von Keyserlingk et al. 2012). Lameness is associated with reduced milk yield (Warnick et al. 2001; Green et al. 2002; Archer et al. 2010), impaired reproductive performance (Garbarino et al. 2004; Walker et al. 2008, 2010), increased risk of culling (Barkema et al. 1994; Booth et al. 2004), and impaired animal welfare (Nordlund et al. 2004; Rushen et al. 2007). These effects result generally in increased production costs (Bruijnis et al. 2010; Cha et al. 2010).

Lameness is commonly detected with locomotion scoring methods. Locomotion scoring can be done quickly on-site, requires no technical equipment, and can be applied easily to a large number of animals (Whay 2002; Flower & Weary 2009; Ito et al. 2010). On the other hand, locomotion scoring is sensitive to variation between and within raters (Engel et al. 2003; O’Callaghan et al. 2003; Thomson et al. 2008; Channon et al. 2009). The quality of subjective measurements is commonly expressed by calculating intra- and inter-rater reliability and agreement (Martin & Bateson 1993; Kottner et al. 2011). Reliability is defined as the capability of raters using locomotion scores to differentiate among individuals (Kottner et al. 2011), eg capability to differentiate between cows scored in level 1 and level 2. Agreement indicates the capability of raters to assign identical locomotion scores to an individual (Kottner et al. 2011).

Lameness is considered a major problem in dairy production.