Observing lame sheep: evaluating test agreement between group-level and individual animal methods of assessment

CJ Phythian*†‡, PC Cripps§, D Grove-White§, PH Jones‡, E Michalopoulou§ and JS Duncan§

* Animal Welfare and Behaviour Group, School of Veterinary Science, University of Bristol, Langford, Bristol BS40 5DU, UK
‡ Department of Epidemiology, University of Liverpool, Institute of Global Health and Infection, and Population Health, Leahurst, Neston CH64 7TE, UK
§ School of Veterinary Science, University of Liverpool, Leahurst, Neston CH64 7TE, UK

* Contact for correspondence and requests for reprints: C.J.Phythian@bristol.ac.uk

Abstract

For on-farm sheep welfare assessment, a reliable, simple and robust method is required to assess the level of flock lameness. This study examined the level of test agreement for two binary lameness scoring systems for sheep. The first was a group-level lameness assessment of sheep performed on ungathered sheep at pasture and was termed group observation method (GOM). The second method of lameness assessment was performed after gathering of the sheep and involved close observation of the gait of individual sheep in a handling pen and was termed individual animal gait assessment (IAGA). Following individual gait assessment, each sheep was also examined for the presence of specific foot and limb lesions: white line lesions (WL); inter-digital dermatitis (ID); footrot (FR); contagious digital dermatitis (CODD); toe granuloma (TG); and joint swellings (JS). A total of 3,074 sheep were assessed from 40 flocks in North England and Wales by one assessor. Test agreement between the assessment methods was found to be good as judged by linear regression and Bland-Altman plots. The method of group observation identified a slightly higher proportion of lame sheep compared to the individual animal examination and also appeared to be a more feasible on-farm method of observation. Over half of the sample sheep were identified with WL but this did not appear to be associated with a high level of lameness (as assessed by IAGA) with just under 12% of sheep with WL being identified as lame. In contrast, the percentage of lame sheep was most closely associated with CODD and over 80% of animals with this lesion were scored as lame.

Keywords: animal welfare, foot lesion, lameness, on-farm assessment, sheep, test agreement

Introduction

Lameness in sheep is an extremely important welfare and economic issue for the sheep industry (Nieuwhof & Bishop 2005; FAWC 2011). In order to inform welfare assessments and intervention strategies, farmers, veterinarians and welfare inspectors require valid and feasible means of measuring lameness in sheep flocks. The level of lameness within the flock is often assessed by farmers and others by observing the behaviour of groups of sheep at pasture. Alternatively, sheep are collected and handled to facilitate a closer inspection of gait and/or a physical foot examination of individual animals (Hodgkinson 2010). There can be differences in the interpretation and assessment of sheep lameness by veterinary surgeons and farmers (Harkins 2005; Kaler & Green 2008a,b), although more recent work suggests that when clear scoring criteria are provided, better levels of agreement between assessors are achieved (King & Green 2011a; Phythian et al 2012).

Clinical observations, such as the assessment of lameness in sheep, are diagnostic tests. Accordingly, the principles used to validate diagnostic tests are used to assess the test validity. For diagnostic test evaluation, validity is defined as the ability of a test to produce correct test results compared to a reference or ‘gold standard’ (Greiner & Gardner 2000). Since there is no current gold standard for lameness scoring in sheep, validity can only be assessed indirectly. Convergent validity concerns the correlation between tests for the same condition, for example the correlation between different methods of lameness assessment.

Reliability is also an important aspect of test validity (Abramson & Abramson 2008). Therefore, the between- and within-observer reliability of both lameness scoring systems for sheep under examination here have previously been determined (Phythian 2011; Phythian et al 2012).

As well as being valid and reliable, a diagnostic test must be feasible for use and applicable under different conditions (Knierim & Winckler 2009). Since the management and environmental conditions for assessment may vary between flocks it may not be feasible to apply a complex multiple category lameness scoring system under all farm systems or terrains. Therefore, a simplified binary scoring system that allows the assessor to walk amongst groups of sheep and record whether...