The Veterinary Journal 216 (2016) 189–195

Contents lists available at ScienceDirect

The Veterinary Journal

journal homepage: www.elsevier.com/locate/tvjl

Inter-observer agreement for clinical examinations of foot lesions of sheep

C.J. Phythian [a,b,*], P.J. Cripps [a], D. Grove-White [a], E. Michalopoulou [a], J.S. Duncan [a]

[a] Department of Epidemiology, Institute of Global Health and Infection, and Population Health, University of Liverpool, Leahurst, Neston CH64 7TE, UK
[b] Section for Small Ruminant Research, Norwegian University of Life Sciences (NMBU), Høyland, 4325 Sandnes, Norway

ARTICLE INFO

Article history:
Accepted 12 August 2016

Keywords:
Foot lesions
Sheep
Clinical diagnosis
Observer agreement

ABSTRACT

In sheep, the diagnosis of foot lesions is routinely based on physical examination of the hoof. Correct diagnosis is important for the effective treatment, prevention and control of both infectious and non-infectious causes of lameness. Therefore, the aim of this study was to evaluate the level of inter-observer agreement for clinical examination of ovine foot lesions. Eight observers of varying experience, training and occupation performed foot examinations on a total of 1158 sheep from 38 farms across North England and Wales. On each farm, a group of two to four observers independently examined a sample of 24 to 30 sheep to diagnose the presence or absence of specific foot lesions including white line lesions (WL), contagious ovine digital dermatitis (CODD), footrot (FR), inter-digital dermatitis (ID) and toe granuloma (TG). The inter-observer agreement of foot lesion assessments was examined using Fleiss kappa (κ), and Cohen's κ examined the paired agreement between the test standard observer (TSO) and each observer. Scoring differences with the TSO were examined as the percentage of scoring errors and assessed for evidence of systematic scoring bias. With the exception of WL (maximum error rate 33.3%), few scoring differences with the TSO occurred (maximum error rate 3.3%). This suggests that observers can achieve good levels of reliability when diagnosing most of the commonly observed foot conditions associated with lameness in sheep.

Introduction

Lameness is a significant and serious global issue for sheep because of the pain, discomfort and debilitation caused (Welsh et al., 1993; Ley et al., 1995; Fitzpatrick et al., 2006). Research has identified that, globally, footrot is the most common cause of lameness in sheep (Egerton et al., 1989; König et al., 2011). Consequently, a variety of strategies for control and elimination of footrot have been devised. These include control approaches based on the administration of systemic antibiotic treatments and culling of persistently-infected cases (Wassink et al., 2010), and elimination strategies based on prophylactic vaccination and whole-flock culling programs (Egerton et al., 2002, 2004; Gurung et al., 2006).

Whilst footrot may be a common cause of lameness (Kaler and Green, 2008a), clearly not all lameness in sheep can be attributed to the condition. Contagious ovine digital dermatitis (CODD), which results in severe lameness and loss of the hoof capsule, currently presents a serious welfare concern for sheep in the UK (Winter, 2008). To date, there is limited knowledge on the epidemiology of this disease and by comparison with footrot only a few recent trials have examined the efficacy of systemic treatments (Duncan et al., 2011, 2012). In addition, there are a number of other foot conditions, including separation and impaction of the white line of the hoof, toe granulomas, interdigital-hyperplasia, septic- and osteoarthritis, which can also result in gait abnormalities of sheep (Winter, 2004; Winter and Arsenos, 2009; Hodgkinson, 2010). Whilst infectious foot lesions remain the most important concern for flock welfare, it has been suggested that these other hoof lesions, such as separation and impaction of the white line (also known colloquially as 'shelly hoof'), are underreported due to misdiagnosis and confusion with footrot cases (Conington et al., 2010a). This is of great importance since the treatment and control points that are deemed to be effective for one foot condition may not be relevant or appropriate for the control of another lesion or infectious cause. The correct identification of a lesion or disease is essential not only for animal welfare reasons but also for economic considerations in order to assess both the scale and economic impact of the disease. Hence, the ability to correctly diagnose foot lesions is vital for implementing prompt and effective treatments and the long-term prevention and control of lameness in sheep flocks (Kaler and Green, 2008a, 2008b).

The ease and accuracy of using diagnoses based on the clinical appearance of lesions need to be further considered given that there

* Corresponding author.
E-mail address: clare.phythian@nmbu.no (C.J. Phythian).

http://dx.doi.org/10.1016/j.tvjl.2016.08.005

is considerable variation in the visual appearance of ovine foot lesions (Kaler and Green, 2008a). Furthermore, there are recognised differences in the interpretation and assessment of different foot lesions amongst differing assessors, such as veterinary surgeons, farmers and researchers (Kaler and Green, 2008b). Microbiological culture (Pitman et al., 1994) and PCR-based techniques (Moore et al., 2005; Frosth et al., 2012) can be employed to complement clinical examination in the diagnosis of some hoof pathologies. However, the time and financial cost of such methods preclude their routine use. Thus, clinical examination by the producer or a veterinary surgeon remains the mainstay for diagnosis of foot conditions in sheep. Consequently, the practical experience and training of farm professionals and veterinarians in the recognition and correct diagnosis of common foot lesions of sheep is an area that warrants further attention.

The diagnostic abilities of different observers can be examined in terms of the level of inter-observer agreement or reliability. The reliability of both binary and categorical scoring measures can be evaluated using agreement analysis methods such as the kappa coefficient (κ) (Kaler et al., 2009). The agreement analysis presents the degree of observed agreement compared to the agreement expected by chance (Feinstein and Cicchetti, 1990) and has been widely used in veterinary research applications, for example to assess the observer reliability for equine health and welfare indicator assessments (Burn et al., 2009) or lameness scores of sheep (Kaler et al., 2009). The type of κ selected depends on the number of observers involved. Fleiss's κ determines the reliability of multiple observers ($n > 2$) (Fleiss et al., 2003), whereas Cohen's κ (Cohen, 1960) examines the reliability of paired assessments ($n = 2$) such as the level of agreement between a study observer and a reference observer, such as the trainer (Burn et al., 2009). κ can also be used to assess the level of agreement between each study observer and a reference observer, such as the trainer (Burn et al., 2009). Several categorical systems for scoring ovine hoof health conditions, and specifically footrot, have been developed and tested (Egerton and Roberts, 1971; Raadsma et al., 1994; Conington et al., 2008; Foddai et al., 2012). However, for routine on-farm assessments as conducted by producers and veterinarians it may not be necessary to use such detailed scoring systems since a binary scale (presence or absence) could provide sufficient information.

The objective of this study was to examine the level of inter-observer agreement for specific ovine foot lesion conditions, using κ agreement analysis statistics and percentage error rate results.

Materials and methods

Study population

The investigation was a cross-sectional study in which 38 farms, located in a 120 mile radius of the University of Liverpool, School of Veterinary Science, Leahurst were recruited through contact with their local veterinary practice. Once the informed consent of farmers was obtained, each farm was requested to gather a sample of approximately 100 sheep for assessment during July to November 2008. On the day of assessment, each sheep was then assigned a numeric identifier in the order they entered the assessment pen and on each farm 30 sheep were selected for examination using a pre-determined random number system.

Observer population

A pool of eight observers was recruited from the University of Liverpool, School of Veterinary Science comprising undergraduate veterinary and animal science students ($n = 3$) and veterinary surgeons ($n = 5$). Observers were classified as 'experienced' if they had undertaken clinical examinations and foot lesion diagnosis of sheep in the previous year (Table 1), those that did not meet these criteria were classified as inexperienced. On the basis of their experience and role in the design and conduct of the study, observer 1 was designated the 'test standard observer' (TSO) and used as the reference test for comparison. All observers were provided with a scoring definition for each lesion, which they were requested to familiarise themselves with together with example images of the specific lesions. In addition, observers classed as 'trained' ($n = 5$) attended a 1-day on-farm training session at the University of Liverpool sheep farm in the diagnosis of foot lesions in sheep. The TSO performed assessments on all study farms and was accompanied at each assessment visit by one to two observers who performed independent clinical examinations of the same sheep on the same day.

Table 1
Description of the observer population.

Observer	Training	Experience	Occupation
1	Trainer	Experienced	Veterinary surgeon
2	Trained	Inexperienced	Veterinary surgeon
3	Trained	Inexperienced	Animal science student
4	Trained	Inexperienced	Veterinary science student
5	Untrained	Inexperienced	Veterinary surgeon
6	Untrained	Experienced	Veterinary surgeon
7	Trained	Inexperienced	Animal science student
8	Trained	Experienced	Veterinary surgeon

Foot examination

Each observer independently performed a clinical examination of each foot of all sample animals as described by Hodgkinson (2010). The absence or presence of any foot lesion in each sheep was recorded. The following specific diagnoses were made based on the descriptions of Winter (2004): white line lesion (WL) – separation and detachment of the white line ('shelly hoof') with impaction or infection present: inter-digital dermatitis (ID) – a raw to white, moist hairless area, progressing to inflammation, infection and necrosis of the inter-digital skin: footrot (FR) – separation of the horn of the hoof, beginning at the junction of the skin and horn near the heel, through to invasion of the sole with separation of insensitive and sensitive laminae: contagious ovine digital dermatitis (CODD) – ulceration around the coronary band, with or without loosening of the claw through to the total loss of the hoof capsule and presence of a raw stump of sensitive laminae: toe granuloma (TG) – strawberry-like growth of proud flesh, which may be covered with loose horn: inter-digital hyperplasia (IH) – folds or protrusions of the skin of variable size located within the inter-digital cleft, and pedal joint sepsis (PJS) – presence of heat, swelling and hair loss above the coronary band, with or without discharging tracts of pus above the coronary band or interdigital cleft. No diagnosis was recorded if it was not possible to make a specific diagnosis based solely on the visual appearance of the foot. Each observer manually recorded their findings on pre-tested recording charts. Observers were not provided with any clinical or production information before each visit. During the visit, each study observer performed an independent clinical examination and observers did not disclose or discuss their foot scores at any stage. The study was approved by the University of Liverpool Ethics Committee (RETH000287).

Data analysis

Data was analysed using Minitab version 16 and Stata version 13 (StataCorp LP). The prevalence (percentage) and 95 percent confidence interval (95% CI) of each foot condition was determined as the total number of sheep observed by the TSO with each foot condition divided by the total number of sheep assessed.

The overall level of inter-observer reliability of multiple observer assessments ($n \geq 2$) was determined by Fleiss's κ (Fleiss et al., 2003). As Fleiss's κ analysis provides an overall agreement value and does not take account of observer characteristics i.e. 'experienced' versus 'inexperienced' assessors, the paired agreement between the TSO and each observer was estimated using Cohen's κ statistic (Cohen, 1960). All κ results were interpreted according to Fleiss et al. (2003), whereby values ≥ 0.75 suggested 'excellent', κ 0.40–0.75 indicated 'fair–good', and $\kappa \leq 0.40$ suggested 'poor' levels of agreement.

As the κ analytical approach cannot identify whether systematic scoring differences occur between pairs or groups of multiple observers, additional approaches were employed to assess the level of observer disagreement in terms of scoring divergence from the TSO. Firstly, scoring differences between the TSO and each observer (TSO score minus observer score) were graphically represented and visually examined for evidence of systematic scoring bias i.e. if an observer consistently scored one unit higher or lower than the TSO. For each observer, the total number of lesions diagnosed by the TSO during paired assessments was calculated and the number of paired scoring differences with the TSO divided by the total number of sheep examined was expressed as a percentage (percentage error rate). Secondly, the proportion of scoring differences with the TSO on each farm visit was plotted to assess if there was any effect of increasing experience of foot assessments on the amount of scoring disagreements. Observers were not provided with any clinical or production information before each visit. During the visit, each study observer performed an independent clinical examination and observers did not disclose or discuss scores at any stage.