The reliability of detecting digital dermatitis in the milking parlour

J.E. Stokes *, K.A. Leach, D.C.J. Main, H.R. Whay

University of Bristol, School of Veterinary Sciences, Langford, Bristol BS40 5DU, UK

A R T I C L E   I N F O

Keywords:
Digital dermatitis
Lesion detection
Milking parlour
Borescope
Dairy cattle

A B S T R A C T

Digital dermatitis (DD) is currently the most problematic infectious skin disease in dairy cattle associated with lameness. Reducing the disease prevalence through early detection and treatment is an essential management tool. The traditional detection method involves lifting and inspecting the feet in a cattle crush, but this is a time intensive and costly practice and impractical for regular detection of individual cases or monitoring herd prevalence. This study aimed to establish the accuracy of detecting and classifying DD lesions in traditional (pit) milking parlours compared with a borescope, and a gold standard lifted foot inspection.

With the exception of one lesion, parlour screening was as accurate as the lifted foot inspection in determining the presence of 86 DD lesions on 160 hind feet (99% agreement; κ 0.99; sensitivity 1.00; specificity 0.99). Describing lesions by colour, depth or stage of lesion in the parlour or using the borescope reached substantial agreement with the gold standard. The stage of lesion was closely linked to colour and depth descriptors. There was greater agreement when categorising more advanced stages of disease progression. Borescope and parlour inspections led to both over and under recording of actual size, particularly in smaller lesions. Screening cows in traditional milking parlours for the presence of DD was found to be an accurate and practical means of detecting lesions. This method should be considered for on farm use to evaluate DD prevention and treatment strategies.

© 2012 Elsevier Ltd. All rights reserved.

Introduction

Digital dermatitis (DD) is a dynamic infectious skin disease of dairy cattle which can be associated with lameness. Reducing the number of individual cases of DD through early detection and treatment, and preventing new cases from occurring are crucial disease management strategies (Döpfer et al., 2012). Traditional inspection of the lifted foot using a crush is logistically challenging as a routine method of detection, particularly as herd sizes increase. Locomotion scoring can be used to identify foot lesions but DD can be present in the absence of lameness (Laven and Proven, 2000).

Rodriguez-Lainz et al. (1998) observed individual cows in a milking parlour for approximately 2 min each, and found an apparent DD prevalence of 20.5% (24/117) compared with an actual prevalence determined by inspection in the crush of 27% (32/117). The sensitivity of milking parlour observation was 0.72, with a specificity of 0.99. However, the sensitivity and specificity may not be a genuine reflection of agreement as the crush examination took place up to 1 month after inspection in the parlour during which time lesion status could have changed. More recently, Thomsen et al. (2008) investigated a rapid screening method taking approximately 15 s to observe each cow in the parlour. The study found a sensitivity of 0.65 and specificity of 0.84 in the parlour compared to inspection in the crush. Taken together these studies suggest that by increasing the time taken to observe each cow, the reliability of detecting DD in the parlour can be increased.

Laven (1999) validated the borescope as a tool for identifying DD in standing cows and reported an individual lesion specificity of 0.84, and sensitivity of 0.82. The technique was subsequently adopted in several research contexts (Laven and Proven, 2000; Vink, 2006; Logue et al., 2012). However, the reliability of a straightforward visual inspection in the milking parlour has not been compared using the borescope. In practice, the expensive and relatively cumbersome nature of the borescope prohibits wide scale application of this technique as a routine detection aid for farm use outside of a research project.

In order to identify a measure that can be used to classify lesions by stage of infection over time, the reliability of different scoring systems has been assessed. A range of characteristics has been used previously to monitor lesion progression (Döpfer et al., 1997; Cruz et al., 2000; Vink, 2006; Holzhauer et al., 2007). Before a scoring system that described lesions by stage of infection was developed, researchers used size, depth and colour (Rodriguez-Lainz et al., 1998; Laven, 1999). These characteristics have the advantage of being simple to measure and easy to apply in a farm context but they do not on their own depict the stage of infection.
Scoring systems which describe the various stages of infection were developed by Döpfer (1994) and Vink (2006). Although there is much agreement between these systems, Vink (2006) is simpler to explain to farmers and apply practically and was therefore used in the current study.

The aim of our study was to identify a practical and efficient means of monitoring the prevalence and severity of DD for a whole milking herd on a regular basis. The reliability of detecting and classifying DD in the hind feet of cows standing in milking parlours that incorporated a pit to allow an eye level view of cows’ feet was compared to the previously validated methods of visualisation using a borescope and lifting the hind feet in a crush (Laven, 1999; Laven and Proven, 2000). The three techniques were evaluated in terms of their sensitivity and specificity of detection, and the agreement achieved when assessing the colour, depth, size (Laven, 1999) and stage of lesion (Vink, 2006).

Materials and methods

Farm details

The study was carried out on three farms with Holstein Friesian cattle in the South-west of England. The herd sizes were 120, 140 and 200 cows, respectively. Herds were selected if they were endemically infected with DD and the herdsmen was willing for a researcher to examine the cows during an afternoon milking followed by an inspection in the farm crush the following morning. Visits to the three farms were carried out each week over a period of 12 weeks. All farms had herring-bone parlours to enable inspection from the pit and a farm crush for the lifted foot examination. Data were collected weekly between mid-February and mid-May, 2008.

Cow selection

On each weekly visit, between two and four cows not previously recruited to the study were selected during milking to fit into one of two groups, namely, (1) one to two cows with no lesions on the hind feet (DD negative), and (2) one to two cows with a DD lesion of varying size (up to 6 cm in diameter at the widest point) on the hind feet. Cows were selected for the DD negative group if they had no visible lesions at the heel, coronary band or interdigital space, and those for the DD positive group if they had no other skin disease visible in the standing foot at the heel, coronary band or interdigital space.

A sampling strategy was employed where one cow per row was inspected. Prior to cluster attachment, and starting at the end of the row, each cow’s hind feet were cleaned with a medium pressure hose and inspected until a cow eligible for one of the two groups was identified. The nozzle was angled downwards across the hoof with water hitting the heel at an approximately 45° angle to ensure minimal splash up. The plantar aspect of each foot at the pastern was carefully examined for DD lesions with the aid of a head torch. Digital dermatitis lesions were scored according to several lesion descriptors. All scoring was carried out by the first author.

Inspection in the crush

Following milking the next morning, cows selected at the previous milking were moved into the farm crush. The hind feet up to the dew claws were cleaned with a medium pressure hose and then dried with a paper towel. While the animal was still standing each hind foot was inspected using the borescope (Fig. 1) and scored according to the lesion descriptors. The crush was then used to lift each hind foot in turn; the feet were then visually inspected and scored using the lesion descriptors. All active DD lesions were treated with oxytetracycline spray. The claws of each hind foot were trimmed using the functional foot trimming method.

Lesion descriptors

The range of lesion descriptors were: stage of infection described by Vink (2006) (Fig. 2), depth (Fig. 3) and colour (cream/yellow, grey/brown/black, or pink/red) assessed visually according to Laven (1999) and size (greatest diameter) measured in millimetres with a tape measure. The size was measured in the standing foot in the parlour and using the borescope by positioning the tape measure as close to the lesion as safely possible. When different stages and colours were present within a single lesion, the description covering the largest surface area was recorded.

Materials and methods

Farm details

The study was carried out on three farms with Holstein Friesian cattle in the South-west of England. The herd sizes were 120, 140 and 200 cows, respectively. Herds were selected if they were endemically infected with DD and the herdsmen was willing for a researcher to examine the cows during an afternoon milking followed by an inspection in the farm crush the following morning. Visits to the three farms were carried out each week over a period of 12 weeks. All farms had herring-bone parlours to enable inspection from the pit and a farm crush for the lifted foot examination. Data were collected weekly between mid-February and mid-May, 2008.

Cow selection

On each weekly visit, between two and four cows not previously recruited to the study were selected during milking to fit into one of two groups, namely, (1) one to two cows with no lesions on the hind feet (DD negative), and (2) one to two cows with a DD lesion of varying size (up to 6 cm in diameter at the widest point) on the hind feet. Cows were selected for the DD negative group if they had no visible lesions at the heel, coronary band or interdigital space, and those for the DD positive group if they had no other skin disease visible in the standing foot at the heel, coronary band or interdigital space.

A sampling strategy was employed where one cow per row was inspected. Prior to cluster attachment, and starting at the end of the row, each cow’s hind feet were cleaned with a medium pressure hose and inspected until a cow eligible for one of the two groups was identified. The nozzle was angled downwards across the hoof with water hitting the heel at an approximately 45° angle to ensure minimal splash up. The plantar aspect of each foot at the pastern was carefully examined for DD lesions with the aid of a head torch. Digital dermatitis lesions were scored according to several lesion descriptors. All scoring was carried out by the first author.

Inspection in the crush

Following milking the next morning, cows selected at the previous milking were moved into the farm crush. The hind feet up to the dew claws were cleaned with a medium pressure hose and then dried with a paper towel. While the animal was still standing each hind foot was inspected using the borescope (Fig. 1) and scored according to the lesion descriptors. The crush was then used to lift each hind foot in turn; the feet were then visually inspected and scored using the lesion descriptors. All active DD lesions were treated with oxytetracycline spray. The claws of each hind foot were trimmed using the functional foot trimming method.

Lesion descriptors

The range of lesion descriptors were: stage of infection described by Vink (2006) (Fig. 2), depth (Fig. 3) and colour (cream/yellow, grey/brown/black, or pink/red) assessed visually according to Laven (1999) and size (greatest diameter) measured in millimetres with a tape measure. The size was measured in the standing foot in the parlour and using the borescope by positioning the tape measure as close to the lesion as safely possible. When different stages and colours were present within a single lesion, the description covering the largest surface area was recorded.
The stage of lesion was closely linked to lesion colour. Consequently, pink/red (erosive) lesions were underestimated (5.8% and 7%, respectively) compared with 16.3% identified during the lifted foot examination (Table 1). Conversely, cream/yellow (granulomatous) lesions were overestimated during the parlour and borescope inspections (77% and 71%, respectively) compared with the lifted foot examination (61.6%).

Measuring the diameter at the widest part of the lesion in the parlour and using the borescope inspection did not meet perfect agreement with the measurement taken in the crush, since parlour and borescope inspections led to both over and under recording of actual size. The greatest inaccuracies were associated with measurement of smaller lesions.

### Inferential statistics

Tables 2–5 summarise the percentage agreement, \( \kappa \) coefficient, sensitivity and specificity results between the parlour, borescope and examination in the crush for each binary parameter at a foot level. All \( \kappa \) values were significantly different from 0 \( (P < 0.001) \). As the agreement between the parlour, borescope and examination in the crush for the presence and absence of DD was near perfect, all feet with a DD lesion regardless of lesion size were analysed together.

With the exception of one lesion, the parlour screening (99% agreement, \( \kappa \) 0.99, sensitivity 1.00, and specificity 0.99) was as accurate as the crush in determining the presence of DD lesions. Classifying lesions by colour (\( \kappa \) 0.61), stage of infection (\( \kappa \) 0.68), or depth (\( \kappa \) 0.63) in the parlour only met substantial agreement with the lifted foot inspection (Table 5).

### Discussion

The aim of this study was to find a practical and accurate method for regular DD detection. With the exception of one lesion, the parlour screening was as accurate as lifting the foot in the crush in determining the presence of DD lesions. The borescope was found to have no added value for screening cows. Screening cows in parlours that have a milking pit was an accurate and practical means of detecting DD.

The cow sample used here was the size of a small milking herd (80 cows), and no lesions in the interdigital space were present. Therefore, the reliability of detecting lesions in the interdigital space in the parlour or using the borescope could not be assessed.
Table 2
The percentage agreement, \( \kappa \) values, sensitivity and specificity for the presence and absence of digital dermatitis and the colour of lesions.

| Detection method       | Presence/absence of lesion | Colour of lesion | | |
|------------------------|----------------------------|------------------|------------------|------------------|------------------|------------------|------------------|
|                        | Percentage agreement       | Kappa            | Sensitivity      | Specificity      | Percentage agreement       | Kappa            | Sensitivity      | Specificity      |
| Parlour and crush      | 99                         | 0.99             | 1.00             | 0.99             | 0.38                 | 0.36             | 0.99             | 0.68             | 0.84             | 0.95             | 0.68             | 0.77             | 0.95             | 0.67             | 0.98             |
| Borescope and crush    | 100                        | 1.00             | 1.00             | 1.00             | 0.59                 | 0.36             | 0.93             | 0.70             | 0.86             | 0.87             | 0.77             | 0.54             | 0.89             | 0.78             | 0.90             |
| Parlour and borescope  | 99                         | 0.99             |                  |                  | 0.46                 | 0.36             | 0.93             | 0.70             | 0.86             | 0.87             | 0.77             | 0.73             | 0.93             |                  |                  |

Table 3
The percentage agreement, \( \kappa \) values, sensitivity and specificity for the stage of lesion.

<table>
<thead>
<tr>
<th>Detection method</th>
<th>Erosive</th>
<th>Granulomatous</th>
<th>Proliferative</th>
<th>Regressed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kappa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Percentage agreement</td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Percentage agreement</td>
</tr>
<tr>
<td>Parlour and crush</td>
<td>0.64</td>
<td>0.55</td>
<td>0.99</td>
<td>0.55</td>
</tr>
<tr>
<td>Borescope and crush</td>
<td>0.65</td>
<td>0.65</td>
<td>0.94</td>
<td>0.70</td>
</tr>
<tr>
<td>Parlour and borescope</td>
<td>0.78</td>
<td>0.67</td>
<td>0.85</td>
<td></td>
</tr>
</tbody>
</table>

Table 4
The percentage agreement, \( \kappa \) values, sensitivity and specificity for the depth of lesions.

| Detection method       | Depth of lesion | | |
|------------------------|-----------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
|                        | Surface         | Shallow          | Deep             | Protruding       |
|                        | Kappa           | Percentage       | Sensitivity      | Specificity      | Kappa           | Percentage       | Sensitivity      | Specificity      | Kappa           | Percentage       | Sensitivity      | Specificity      | Kappa           | Percentage       | Sensitivity      | Specificity      |
| Parlour and crush      | 0.57            | 0.52             | 0.79             | 0.73             | 0.79             | 0.92             | 0.57             | 0.96             | 0.71             | 0.92             | 0.63             | 0.97             | 0.86             | 0.98             |
| Borescope and crush    | 0.65            | 0.63             | 0.83             | 0.84             | 0.85             | 0.96             | 0.71             | 0.92             | 0.71             | 0.92             | 0.74             | 0.98             | 0.71             | 0.95             |
| Parlour and borescope  | 0.68            | 0.72             | 0.87             | 0.72             | 0.76             | 0.95             |                   |                  | 0.77             | 0.97             |                  |                  |                  |                  |

Table 5
The overall percentage agreement, \( \kappa \) values, sensitivity and specificity for each lesion descriptor.

<table>
<thead>
<tr>
<th>Detection method</th>
<th>Size</th>
<th>Colour</th>
<th>Stage</th>
<th>Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kappa</td>
<td>Percentage agreement</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>Parlour and crush</td>
<td>85</td>
<td>0.90</td>
<td>0.61</td>
<td>90</td>
</tr>
<tr>
<td>Borescope and crush</td>
<td>89</td>
<td>0.90</td>
<td>0.63</td>
<td>90</td>
</tr>
<tr>
<td>Parlour and borescope</td>
<td>89</td>
<td>0.90</td>
<td>0.61</td>
<td>90</td>
</tr>
</tbody>
</table>
In describing the stage of infection, the $\kappa$ coefficients calculated between parlour, borescope and lifted foot inspection, for the stage, colour and depth of lesions all met the criterion for substantial agreement. However categorising lesions by stage in the parlour and using the borescope was less accurate than when the lifted foot was inspected. It was more difficult to distinguish between early stages of lesions (erosive and granulomatous) using the borescope and the parlour, compared to classifying later stages of the disease (proliferative and regressing). However such distinctions may have more value in a research context than a disease management setting. As evidenced by the relationships seen between the results in Table 1, the stage scoring system inherently incorporates colour and depth characteristics, as well as describing the development of infection over time.

One purpose of this study was to identify an accurate lesion descriptor for assessing the efficacy of treatment interventions. Lesion stage assessed in the parlour met substantial agreement with the lifted foot inspection. Size (diameter) of lesion met better agreement although both the borescope and parlour assessments led to both an over and under recording of actual size.

The difference in statistical agreement can be partially explained by the methods used to analyse the data. Kendall’s coefficient of concordance which was used to analyse the size (diameter) of lesions took into consideration the degree to which two methods do not agree. However the $\kappa$ statistic used to analyse the lesion stage measure was assessed as a binary classification only, so did not take into account degree of agreement.

Greater agreement was found between the current experimental methods than some previous studies have reported (Rodriguez-Lainz et al., 1998; Thomsen et al., 2008). However, recently Relun et al. (2011) reported a sensitivity of 0.90 and specificity of 0.80 in detecting the presence and absence of lesions in the parlour. This variation can be attributed to differences in experimental method, such as the time taken to observe each cow as well as differences in observer training, experience and accuracy. In the present study the researcher only selected one cow per row, which meant there was approximately 7 min per cow to score a number of lesion characteristics. As a range of lesion characteristics were scored, this study did not try to detect lesions across a whole herd during one milking. It was also beyond the scope of the study to assess time taken to score each cow. Although the time available to us to make observations during our study is unrealistic under real life conditions, it is unlikely that such a wide range of observations would be made together in a practical farm situation. Subsequent to the study described here, this protocol, using a single lesion descriptor has been successfully integrated into the milking routine without slowing down milking. The hind feet of whole herds were scored to determine the presence and stage of infection across fifteen herds ranging from 73 to 305 cows (Stokes, 2011).

One limitation to this study design was that it did not assess inter- and intra-observer agreement, however recently published data (Relun et al., 2011) shows that inter- and intra-observer agreement is relatively high for in-parlour detection of DD. Those authors assessed agreement in 242 cows across four farms. The hind feet were inspected for the presence and stage of DD by five veterinarians during two consecutive milkings using a head torch and swivel mirror. Good reliability was found between assessors and repeatability across milkings for the presence/absence of DD and this agreement was improved when stages of DD (Döpfer et al., 1997) were amalgamated into three stages (absence: M0/acute lesion: M1–M2/chronic lesion: M3–M4) with high intra-observer reliability ($\kappa 0.62 [0.51–0.65]$) and inter-observer reliability ($\kappa 0.63 [0.51–0.65]$).

There was a selection bias inherent in the study design as, for practical reasons, cows were always selected in the milking parlour first and then subsequently re-assessed using the borescope and the lifted foot examination. Also, cows were not randomly selected as a balance of infected and non-infected individuals was necessary to allow identification of both true positives and negatives. In addition it was not possible for the observer to be blinded to the selection history of the cows; therefore the results may be biased by previous examinations. This poses a threat to the internal validity of the study. Efforts were made to reduce the observer effect by carrying out the borescope inspection a day after the parlour inspection, employing a research technician to record information onto the data collection sheet during the borescope and crush assessments, thus reducing the potential for the researcher to review scores allocated to the cows during the parlour screening. As the parlour and subsequent observations were taken a day apart, the parlour inspection may have influenced subsequent scoring during the borescope and lifted foot inspections. However it has to be noted that the borescope and lifted foot inspections were carried out in immediate succession. If the observer bias affected the way lesions were scored, substantially more agreement between the borescope and the lifted foot inspection than with the parlour screening would be expected. This was not found to be the case.

Regularity screening cows for DD in the milking parlour can lead to early treatment, reducing the infection reservoir within a herd which is currently understood to be the lesion itself (Carter et al., 2009; Döpfer et al., 2012). Regular cleaning of feet can make individual DD detection easier and early treatment interventions may reduce the likelihood of developing chronic cases (Somers et al., 2003; Döpfer et al., 2012).

Conclusion

In the case of routine assessments, the protocol to evaluate DD prevention and treatment strategies is largely shaped by the feasibility of regular whole herd lesion monitoring. Detecting lesions in the early stages of infection is important from both welfare and disease management perspectives. Examination of cows in the parlour following the method described here can be used to identify accurately the presence and absence of DD lesions and/or to describe lesions characteristics and is sufficiently robust to be recommended as a method for routine monitoring of DD prevalence and severity on farm.

Conflict of interest statement

None of the authors of this paper has a financial or personal relationship with other people or organisations that could inappropriately influence or bias the content of the paper.

Acknowledgements

This study was supported by Tubney Charitable Trust, Soil Association, Freedom Food, The Organic Milk Suppliers Cooperative, Long Clawson, and Milk Link. Thanks go to the farmers who enabled data collection.

References


