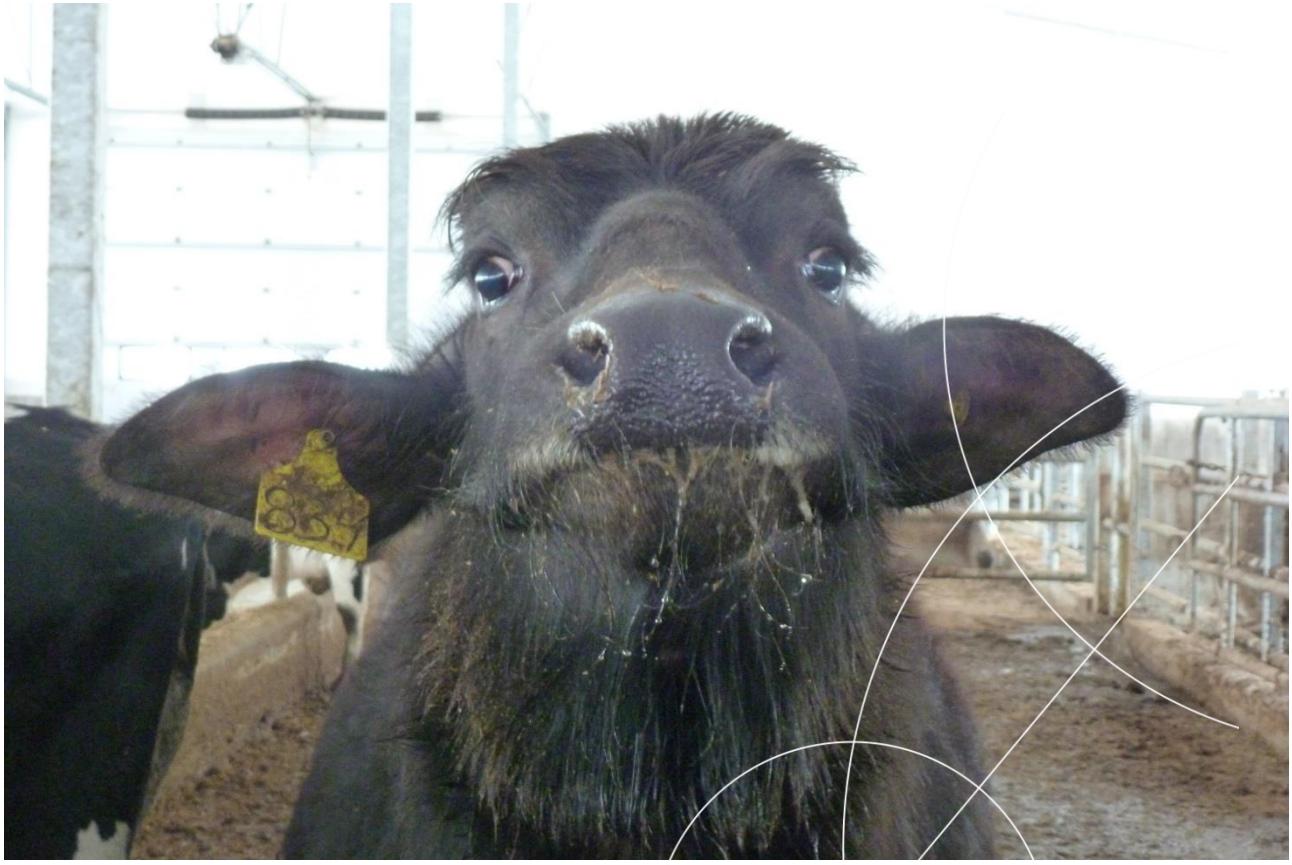
FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Master's thesis

Rita Palmelund Kviesgaard QNF256

Quality of hoof health records

- evaluated with analysis of inter- and intra-observer agreement



Academic advisor: Nynne Capion DVM, Ph.D.

Submitted: 19/04/2013

Institutnavn: Institut for Produktionsdyr og Heste

Name of department: Department of Large Animals Sciences
Højbakkegård Allé 5
2630 Taastrup

Author: Rita Palmelund Kviesgaard
Regensen
St. Kannikestræde 2
DK-1169 København K
Phone no: +45 26842647
Email: ritapalmelund@gmail.com

Master's thesis

Title / Subtitle: Quality of trimming records

- evaluated with analysis of inter- and intra-observer agreement

Subject description: In this observational study the quality of trimming records was evaluated. From paired records between a gold standard and 18 hoof trimmers, the inter observer agreement was calculated and used as a measurement of trimming record quality.

Academic advisor: Nynne Capion DVM, Ph. D.

Submitted: 19. april 2013

Photographer: Rita Palmelund Kviesgaard

(title page picture)

Resume

Vigtige beslutninger i mælkeindustrien bliver taget på baggrund af klovbeskæringsregistreringer fra rutine klovbeskæring. Formålet med denne opgave var at undersøge om uddannelse af klovbeskærer har effekt på kvaliteten af klovbeskæringsregistreringer. Atten tilfældigt udvalgte klovbeskærer deltog i studiet, fem certificerede og otte ikke certificerede klovbeskærer fra Danmark, samt fem trænede canadiske klovbeskærer. Under rutine klovbeskæring blev der indsamlet parrede registreringer fra 9100 klove. Graden af enighed mellem en veterinærstuderende (den gyldne standart) og hver enkelt klovbeskærer for hver lidelsestype, udtrykt som vægtet kappaværdi, blev brugt som et kvalitetsmål for klovbeskæringsregistreringerne.

Resultaterne viste at der var stor spredning i kvaliteten både mellem klovbeskærer og lidelsestyper. Mellem de tre forskellige grupper af klovbeskærer var der ikke betydelig forskel i registreringskvalitet. Der var dog en klar tendens til at balleråd, såleblødninger og hul væg, ikke blev registreret i samme grad som sålesår, digital dermatitis og nydannelser. Dette blev tolket som et tegn på at klovbeskærerne er mindre motiverede for at registrere visse lidelsestyper.

Abstract

Hoof health records based on routine trimming of dairy cows serve as basis for important decisions in the dairy industry. The objective of this observational study was to clarify if education of hoof trimmers had effect on the quality of trimming records. A total of 18 randomly chosen trimmers participated in the survey; five certified, and eight uncertified trimmers from Denmark plus five trained Canadian trimmers. Paired data from 9,100 hooves collected during routine trimming was included in the study. Inter-observer agreement expressed as weighted kappa values was used as a measurement for record quality. It was calculated for each trait between a veterinary student, who was the gold standard, and each trimmer.

The results showed great dispersion in record quality between both trimmers and lesion types. There was no considerable difference in the quality between the three groups of trimmers. However, there was a clear pattern that heel erosion, sole hemorrhages, and white line lesions had a much lower recording rate compared to sole ulcers, digital dermatitis, and interdigital hyperplasia. This was interpreted as trimmers being less motivated for recording certain lesion types.

Keywords: hoof trimmer, trimming record, inter-observer agreement, dairy cow

Introduction

Lameness is a problem in the dairy industry both economically and from an animal welfare point of view (O'callaghan et al., 2003; Ettema et al., 2009). A study from 2008 revealed a high prevalence of hoof lesions in Danish Holstein herds and underlined a need for more information on the subject (Capion et al., 2008b). Hoof health is affected by both management and genetics (Van der Linde et al., 2010; Bergsten, 2010). Collaboration between The Danish Cattle Federation and the cattle breeding company Viking Genetics lead to the development of the computer software 'Klovregistrering', that has the primary objective of having hoof trimmers collect data on foot lesions, which can be used in breeding programs for improving hoof health in Scandinavia (NAV,

2011). In addition, dairy producers benefit from the trimming records when setting up lameness reduction strategies in coordination with veterinarians, industrial partners and researchers (Manske, 2003; Seested and Thomsen, 2011; Daniel, 2012).

Professional Danish hoof trimmers have been making digital records of foot lesions during regular trimming of dairy cattle since 2010. Both certified and uncertified trimmers can upload records to the Danish Cattle Database. The certified trimmers (CTs) have had specialized training and passed a test in lesion identification and severity scoring as part of their education (N. Capion, personal communication). Education is currently not required for hoof trimmers to operate in Denmark. It is presumed that CTs make trimming records of higher quality compared with uncertified trimmers (UTs).

Data on hoof health from the Danish Cattle Database shows great deviation in lesion prevalence in between trimmers, which indicates that the recordings vary greatly depending on the individual trimmer (Nielsen, personal communication). In The Ontario Dairy Hoof Health Project from 2012, Canadian hoof trimmers (TTs) received three days of training in standardized lesion identification and severity scoring before they started operating the chute side touch screen for data recording (Daniel, 2012). By using a gold standard[1], it is possible to compare the quality of trimming records of Danish and Canadian trimmers respectively.

A limited amount of literature on trimming records exists, and more knowledge will aid the evaluation of hoof health records as well as assess the quality of the data for later use. The objective of this study was to evaluate the effect of training of hoof trimmers on the quality of their recordings and propose ways of improving the record quality.

[1] a thing of superior quality which serves as a point of reference against which other things of its type may be compared (Oxford Dictionary)

Materials and methods

Study design

The study was designed as a prospective observational study of the quality of trimming records from dairy cattle made by CTs, UTs, and TTs respectively. Data collection was carried out from October 2012 to March 2013.

Thirteen Danish trimmers were included in the study; five CTs and eight UTs. They were randomly selected from a list of hoof trimmers performing a minimum of 5,000 recordings per year. The study was blinded, as the veterinary student (RPK), who was the gold standard, did not know beforehand which of the trimmers were certified and which were not. In addition, five TTs were included in the survey. Originally 16 Danish trimmers were chosen, but because of a limited time frame and practical reasons, three were excluded from the study.

In advance RPK had received two days of intensive training in lesion recognition and severity scoring by two skilled CTs recommended by DVM. Capion, N.

The trimmers were contacted by phone and asked to participate in the study, and they all accepted. Each trimmer was told that RPK had to observe the trimming of about 150 Holstein cows, and herd hoof health should be average or lower to make sure there were enough lesions to record.

The ability to recognize and lesion score trimmed hooves with common lesions was tested on 20 photographs present to the trimmers, who had to fill out an answer sheet using the trait definitions given in Table 1 and Table 2. The results were compared to the gold standard.

In order to evaluate the reliability of RPK's records, an intra-observer agreement was calculated before and after the data collection. By lesion scoring (using the definitions in Table 1) of 48 pictures of trimmed feet two days in a row, K was calculated for each trait and used as a measurement of consistency.

Data collection

Data sampling took place during regular hoof trimming. The trimmers were told to do as usual, not letting the presence of the observer (RPK) influence on their recordings. They were not told in advance what the objective of the survey was.

Hooves were trimmed with a grinder by the hoof trimmer. RPK observed the trimming act and before a foot was put down or wrapped, a quick inspection was made. To avoid influencing the trimmers' registrations, no further examination of the feet were made. Every single hoof was evaluated; lesions and the severity were noted on paper independently of the trimmer. Multiple lesions of the same type on any hoof counted as one. The trimmers entered lesions on the touch screen after each cow. When two trimmers worked together on the same chute, they counted as one person.

Lesions were identified on the basis of macroscopic observations of the skin and claws as described in Table 1 and Table 2. The software for entering lesions and severity (Klovregistrering for Danish trimmers and Hoof Supervisor for Canadian trimmers) is based on the given definitions of the respective country.

Table 1: Trait definitions applying for Denmark, Finland and Sweden

Lesion Name	Severity	Definition
Heel erosion (HE)	Mild	Superficial loss of heel horn.
	Severe	Moderate to profound erosion of bulbs (fissures or craters), possibly extending to the corium.
Sole hemorrhage (SH)	Mild	Superficial hemorrhages or light discoloration of the horn of sole and/or white line, may disappear at normal trimming.
	Severe	Profound hemorrhages of the horn of sole and/or white line, remains after trimming.
Sole ulcer (SU)	Mild	Penetration through the sole horn exposing fresh corium.
	Severe	Penetration through the sole horn exposing affected corium with necrotic discolored tissue and/or protruding granulation tissue. Painful!
White line lesion (WL)	Mild	Separation in the white line, present after normal trimming.
	Severe	Exposed corium with/without purulent exudation originating from a fissure of the white line. (Abscess of white line)
Digital dermatitis (DD)		Infection of the digital and/or interdigital skin with erosion, typical looking, bleeding ulcerations. Painful!
Interdigital hyperplasia (IH)		Interdigital growth of fibrous tissue
Double sole (DB)		Two or more layers of underrun sole horn

The Klovregistrering software is designed to be operated with very few tabs on the screen when lesion recording. After cow ID has been entered the lesion type and severity can be chosen with one single tab whereupon, the affected feet can be chosen and other lesions entered. The trimmer can choose to record lesions on cow level, rather than per foot. Table 1 is based on the definitions given for the Nordic countries (Nielsen, personal communication)

Table 2: Trait definitions applying for Canada. Severity score is based on appearance after trimming.

Lesion Name	Severity	Definition
Heel Erosion		Score only lesions of severity 3
Sole hemorrhage	1	Light colored blood streaks in sole horn
	2	Darker red or blue areas left in the sole after trimming
	3	Very dark red, purple or blue areas left in the sole after trimming
Sole ulcer	1	The corium is exposed
	2	Exposed granulation tissue is evident
	3	Exposed granulation tissue is larger than the end of a pinky finger
White line lesion	1	Ranges from hemorrhage to slight separation of white line. Trimming may release small amounts of pus
	2	Minor exposure of corium; detachment of horn, substantial pus
	3	More pronounced exposure of corium, possibly extending up to the coronary band
Digital dermatitis	1	A circular lesion above the heel bulb smaller than 26mm (about an inch) in diameter
	2	A lesion between 26mm and 52mm in diameter
	3	A lesion greater than 52mm in diameter
Interdigital hyperplasia	1	Growth of tissue between claws but does not fill interdigital space
	2	Hyperplastic tissue fills interdigital space
	3	Growth of tissue between claws causes claws to spread

When the cow ID has been entered the Hoof Supervisor software forces the trimmer to first select a claw, then the zone with a lesion, subsequent a menu pops up giving the trimmer the option to choose between different lesions relevant for the given zone, finally the severity is scored. Each claw can be given multiple lesions. Tabel 2 is based on the Canadian Lesion Severity Scoring Guide (Greenough and et al., 2011)

Data was registered as categorical (in Denmark; HE, SH, and SU, in Canada; SH, SU, DD, WL, and IH), or binary (in Denmark; DD, IH, and DB, in Canada; HE) traits according to the definitions. Because of a software defect, WL was excluded from the Danish part of the analysis. To delimit the study only the listed lesions were included.

Statistical analysis

A list with all recorded lesions on individual cows in each herd was printed at the end of every trimming session. Inter-observer agreement between the trimmer and the gold standard was calculated as weighted kappa (K) by using an online calculator for every recorded lesion type (Lowry, 2013). In addition, the 95% confidence interval (CI) was noted. The degree of agreement was interpreted according to Landis and Koch (1977), see table 3. The K was used as a measurement of trimming record quality. Agreement was acceptable when $K > 0.41$.

Table 3: Interpretation of kappa values (Landis and Koch, 1977)

Kappa value	Strength of agreement
0-0.2	Poor
0.21-0.4	Fair
0.41-0.6	Moderate
0.61-0.8	Good
0.81-1	Excellent

Results

Records from 2,437 cows (total number of hooves: 9,100) were included in the study. The mean number of recorded feet per trimmer were 555 (CI 511-600) and 375 (CI 341-409) for Danish and Canadian respectively. Of the 110-115 hoof trimmers in Denmark, approximately 50% take part in the digital recording scheme. In contrast, there are 12 in Ontario. A total of 46 cows were excluded due to discrepancies in cow numbers, repeated numbers and missing values. A complete record may consist of records from two-three hooves.

Traits with high prevalence (p) have a higher accuracy than traits with low p. Therefore a high inter-observer agreement is not necessarily equal to a high accuracy. Trimmer no. 1 and 17 have $K_{SU} > 0.81$, but p_{SH} in the given study herds are 0.04 (Table 4). The high K is due to a high agreement of hooves with no lesion, rather than agreement on lesion existence and severity. In contrast, trimmer no. 6 and 7 have very poor K_{SH} despite $p_{SH} \geq 0.62$ (Table 4) the accuracy of these results are very high.

All diagnoses have been given per foot, except from those marked as lesions at cow level (Table 4). If one cow got the general diagnose severe HE by the trimmer and RPK had recorded mild HE on three legs and severe on one leg, the cow counted as agreement on severe HE. That is only the most severe lesion counted. Trimmer no. 1, 2, 3, 5, 6, 9, 11, and 13 represent two trimmers working together.

Table 4: Number of cows (N cows) and recorded feet (N feet) for each trimmer along with lesion prevalence (*p*) and weighted kappa value (*K*) for heel erosion (HE), sole hemorrhage (SH), sole ulcer (SU), digital dermatitis (DD), interdigital hyperplasia (IH), double sole (DB), and white line lesions (WL).

Trimmer no.	N cows	N feet	*p* HE	*K* HE	*p* SH	*K* SH	*p* SU	*K* SU	*p* DD	*K* DD	*p* IH	*K* IH	*p* DB	*K* DB	*p* WL	*K* WL
1	177	509	0.44	0.41[a]	0.55	0.39	0.04	0.86	0.02	0.5	0.12	0.87	0.14	0.52		
2	125	366	0.18		0.48	0.44	0		0.09	0.75	0.1	0.92	0.08	0.66		
3	173	610	0.28	0.09[a]	0.63	0.37[a]	0.02		0.14	0.81	0.12	0.87	0.05	0.73		
4	225	880	0.70		0.6	0.04	0.07	0.41	0.24	0.41	0.09	0.6	0.08	0.23		
5	99	370	0.52		0.63	0.07[a]	0.07	0.17	0.37	0.6	0.3	0.62	0.11	0.04		
6	55	215	0.38	0.19	0.73	0.02	0.13	0.14	0.4	0.55	0.24	0.85	0.18	0.45		
7	197	757	0.49	0.06	0.62	0.05	0.05	0.36	0.05	0.66	0		0.14			
8	146	571	0.42		0.62	0.07	0		0.18	0.86	0.08	0.8	0.01	0.33		
9	149	537	0.54	0.2	0.78	0.41	0.07	0.58	0.42	0.75	0.15	0.84	0.5			
10	201	783	0.67		0.64	0.07	0.06	0.64	0.28	0.49	0.22	0.69	0.06	0.18		
11	179	694	0.61	0.24	0.77	0.4	0.06	0.69	0.04	0.93	0.02	0.5	0.09	0.28		
12	117	461	0.49		0.68	0.19	0.13	0.69	0.12	0.88	0.08	0.55	0.14	0.55		
13	119	471	0.82		0.77	0.03[a]	0.13	0.64	0.34	0.72	0.19	0.83	0.26	0.08		
14	91	359			0.45	0.39	0.01	0.5	0.29	0.8	0.1	0.71			0.74	0.24
15	101	402			0.29		0.01		0.46	0.74	0.26	0.31			0.46	0.09
16	88	349			0.44	0.38	0.06	0.8	0.56	0.66	0.34	0.32			0.22	
17	102	408			0.33	0.51	0.04	0.93	0.2	0.74	0.03				0.13	0.1
18	93	358			0.39	0.59	0.13	0.58	0.41	0.76	0.26	0.67			0.46	0.37

118 [a] *K* is based on cow level, since the trimmer did not record this lesion at hoof level.

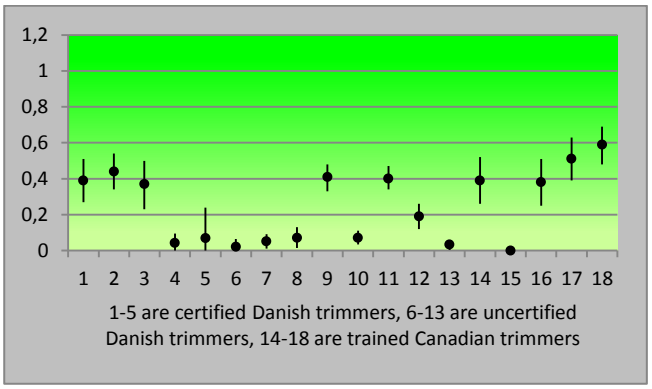


Fig. 1: Weighted kappa values for the agreement between each trimmer and RPK for scoring sole hemorrhages

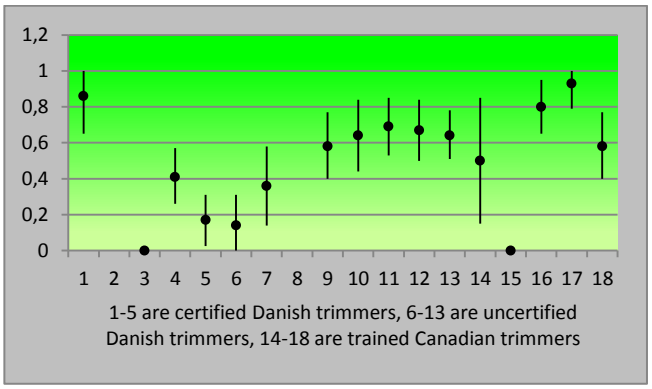


Fig. 2: Weighted kappa values for the agreement between each trimmer and RPK for scoring sole ulcers

119

120 The vertical lines in Fig. 1-Fig. 7 express the 95% CI. The K_{SH} is low as only one CT, one UT, and
121 two TTs obtain moderate agreement. The rest have $K_{SH} \leq 0.4$ and the majority of these display poor
122 agreement (Fig. 1). Relatively, the TTs have the best results for scoring SH. The biggest dispersion
123 in lesion recognition and scoring was found for SU (Fig. 2). Four trimmers have poor agreement
124 while seven out of 16 have $K_{SU} > 0.61$. SU has a low prevalence and the CI indicate that most
125 trimmers can be placed in a different agreement category. The proportion of trimmers with $K_{SU} \geq$
126 0.41 is 50%, 71%, and 80% for CTs, UTs, and TTs respectively (Fig. 8).

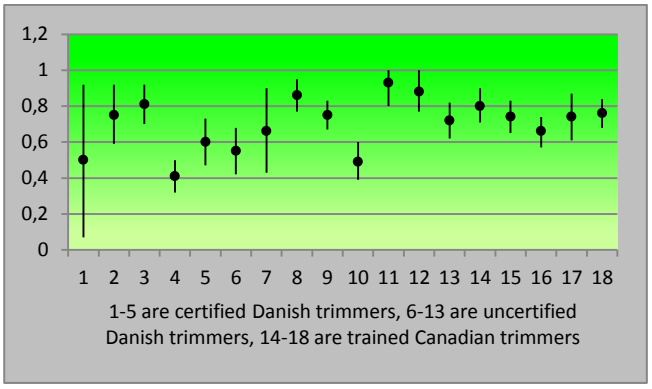


Fig. 3: Weighted kappa values for the agreement between each trimmer and RPK for scoring digital dermatitis

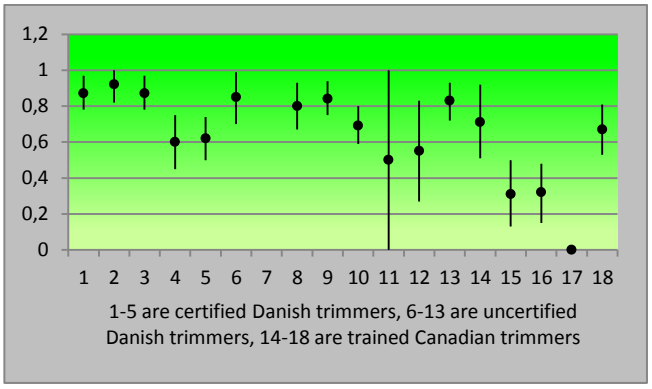


Fig. 4: Weighted kappa values for the agreement between each trimmer and RPK for scoring interdigital hyperplasia

Fig. 3 shows that the trimmers achieved a relatively high K_{DD}. Thirteen out of 18 trimmers had K_{DD} >0.61 with the rest having $0.41 \leq K_{DD} \geq 0.6$. All TTs obtained good agreement, the UTs had two with moderate and three with excellent agreement while the CTs only had two out of five with K_{DD} > 0.6. When scoring IH there is some dispersion and a higher percentage of CTs and UTs having $K \geq 0.61$ compared to the TTs (Fig. 4 and Fig. 9).

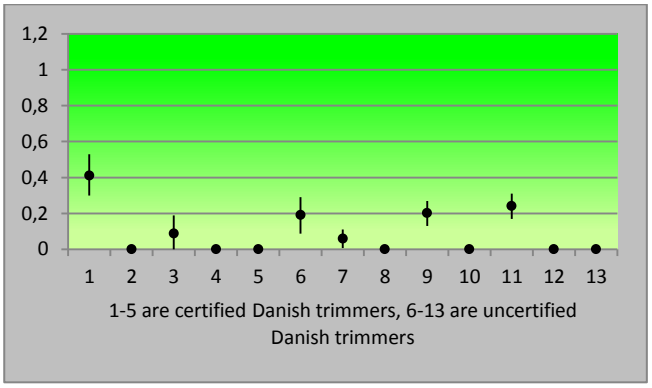


Fig. 5 Weighted kappa values for the agreement between each trimmer and RPK for scoring heel erosion

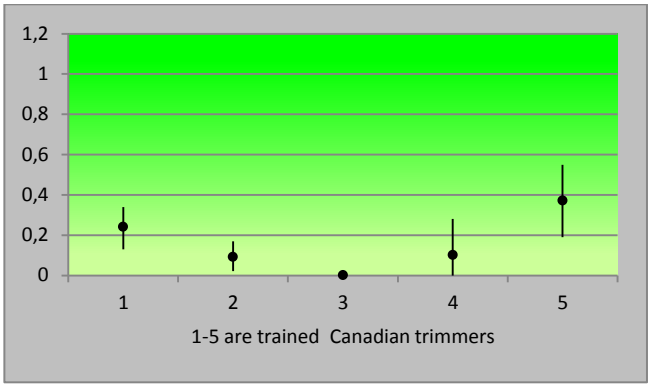


Fig. 6 Weighted kappa values for the agreement between each trimmer and RPK for scoring white line lesions

The K_{HE} has only been computed for Danish trimmers, and all but one CT and one UT have poor agreement (Fig. 5). As Fig. 6 shows, all TTs disagreed highly when scoring WL. Hemorrhages in the white line is defined as a SH in Denmark and a WL in Canada. This can be part of the explanation why CTs and UTs have low K_{SH} and TTs has low K_{WL}.

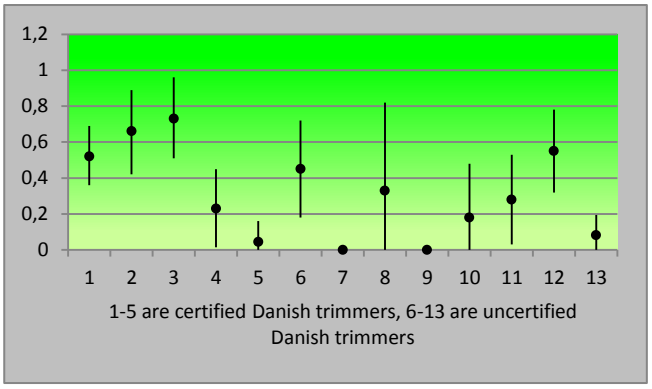


Fig. 7 Weighted kappa values for the agreement between each trimmer and RPK for scoring double sole

Calculated K_{DB} show great variation. Only two UTs and three CTs achieved moderate or good agreement (Fig. 7).

In general, the trimmers had higher agreement when recording SU, DD, and IH compared with HE, SH, and WL. Disagreement was most often due to no recordings of a lesion rather than discrepancy in the severity (data not shown).

Fig. 8 and Fig. 9 illustrate the agreement distribution, which can be used to evaluate the record quality of each group of trimmers. No individual group consistently performs better as it depends on lesion type and cutoff value.

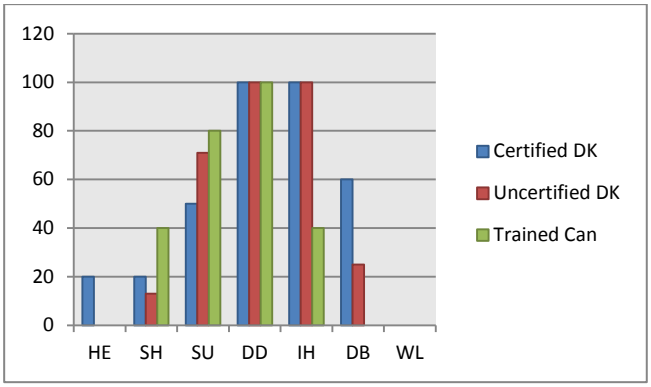


Fig. 8 Percentage distribution of the 18 trimmers with weighted kappa values ≥ 0.4

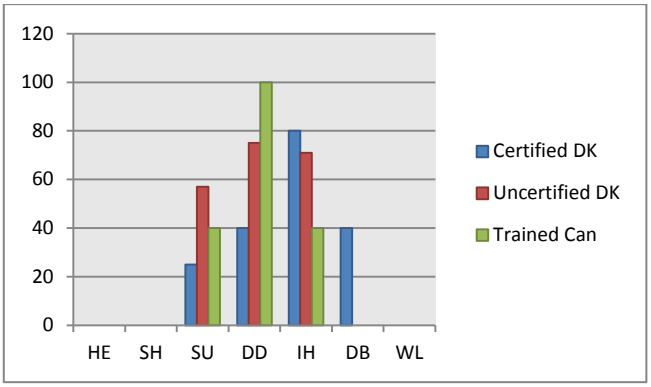


Fig. 9 Percentage distribution of the 18 trimmers with weighted kappa values ≥ 0.6

There is neither a correlation between inter-observer agreement and years of trimming experience, nor time of data collection. Generally, K did not increase or decrease throughout the data sampling period. There was neither consistency in K and CI for individual trimmers, nor for trimmers working in pairs compared with trimmers working alone (data not shown).

The hoof trimmers were asked to lesion score the same 20 photographs and answers were compared with the gold standard. They identified HE (except for TTs), SH, SU, DD, and WL all with

acceptable *K*. The agreement was moderate or better (data not shown) where only one UT had poor K_{HE} and one CT had poor K_{SH} (one UT did not do this test).

RPK's consistency in lesion rating was tested by scoring 48 randomly selected pictures, of trimmed hooves on two successive days before and after data collection. All results of the intra-observer agreement test had K $\geq$0.61 except for K_{SH} before data collection (Table 5). In addition, the reproducibility between before and after was estimated, with results varying from fair to excellent. Before the data collection, the lesions were scored more severely than afterward (data not shown), which is the explanation for the lower agreement.

Table 5: Weighted kappa values for the intra-observer agreement for RPK before and after collection of data, plus agreement between before and after.

	Heel erosion	Sole hemorrhage	Sole ulcer	Digital dermatitis	White line lesion	Interdigital hyperplasia
Before	0.61	0.54	0.88	0.88	0.65	0.62
After	0.81	0.77	1	1	0.67	1
Agreement B-A	0.58	0.34	0.75	0.65	1	0.41

Discussion

Knowledge about trimming record quality is limited. Most often, training sessions have been prior to lesion scoring agreement estimations in other studies (Manske et al., 2002; Capion et al., 2008a; Capion et al., 2008b). In this study, data collection during regular trimming, without interfering with trimmers, was done to get a truthfully picture of the rate of lesion recording under field conditions.

Between Canada and Denmark, some differences regarding trimming records exist (Table 6). In most cases, the differences favored TTs and increased the chance of high agreement.

Table 6: Differences concerning lesion recording between Canada and Denmark

Canada	Denmark
Mandatory introduction to software and severity guide. Trimmers received training in standardized lesion identification and severity scoring both on pictures and live cows. August 2011	Introduction to software and correct scoring of lesions was offered but not required. April 2010
Software decreases risk of data input errors. The foot is divided in 13 zones. For each zone the software lists the lesions relevant for the zone. The same system is used in Germany (Kofler et al., 2013).	Lesions can be entered with a minimum of tabs on the screen. Lesions can be recorded on cow level and does not have to be linked to a single hoof.
Trait definitions are slightly different and most have 3 severities (4 levels) (acc. to table 2).	Some traits are scored as present or not, others as mild or severe (3 levels) (acc. to table 1).
Canadian trimmers only record lesions that are present by the end of trimming; fewer lesions to record.	Danish trimmers are supposed to record even mild lesions that are removed during trimming.
There appeared to be a general acceptance of slower working pace and both the farmer and trimmer seemed more relaxed.	Several Danish farmers encouraged trimmers to go through a whole herd in one day to minimize the disturbance in the herd.
Average no of trimmed cows per day was lower (40-60), which meant less stress and better concentration.	Number of trimmed cows per day often exceeded 100, which might decrease quality of recordings due to fatigue.
Can print out tables after finished trimming session (instant results to show to the farmer and basis for discussions of hoof health situation in herd).	Data is being uploaded to the database, sometimes several days after trimming. Farmer has to download the records to get tables and evaluate hoof health without hoof trimmer.

This table is based on observations, communication with hoof trimmers, Videnscenteret for landbrug, 2010; Greenough and et al., 2011; Daniel, 2012 and Nielsen, personal communication.

It was not possible to do a test to clarify if there was significant difference in the calculated *K* between the three groups of hoof trimmers, because data was collected on different sample populations. However, TTs had slightly higher agreement than CTs and UTs when recording SH, SU, and DD. Table 6 shows examples of circumstances that may explain this.

To increase power of the survey, higher lesion variety and prevalence in study herds were desired, and more hoof trimmers should have been included, but a limited time-frame did not allow for this. A weakness in the study was that RPK's lesion scoring skills were never evaluated by an external professional. Day-to-day repeatability in the intra-observer test was highly satisfactory. However, there was a change in severity scoring.

Several conditions during the data collection could have biased the results. The presence of a veterinary student could have influenced lesion recording, and trimmers calling out lesions to a partner could also have affected RPK. If this was the case, higher agreement could have been expected. To avoid influencing the trimmer, a closer examination of hooves was not performed in

some cases. This has possibly led to overseen lesions (primarily IH and DD) and a lower *K*, but having evaluated the results, this bias has been minimal. Some trimmers recorded lesions according to outdated lesion definitions, which had a negative impact on the agreement. RPK took notes immediately after the trimming of every single hoof, while some trimmers memorized lesions from 2-4 feet before entering data. Lesions that were removed during trimming, such as mild SH and DB, might not have been observed by RPK, as it was not always possible to observe all hooves during trimming, which likely reduced the estimated agreement.

Despite the trait definitions, each foot is a subjective evaluation and personal interpretation of the definitions, which may explain some of the dispersion in *K* between trimmers. The borderline between a mild and severe lesion is different for each individual trimmer. The Canadian lesion severity guide for reference is very detailed with description and several pictures of each lesion severity. All TTs have one, which is used frequently. Inter-observer agreement is generally increased by direct measurements instead of subjective scores (Manske, 2002). The more defined the trait definitions are, the higher the inter-observer agreement ought to be, which was also indicated here as the TTs showed a tendency to perform better under field conditions. An amplified severity guide with more pictures of each lesion and severity may improve the Danish trimmers' recording. Mandatory training is assumed to increase skills in lesion identification and severity scoring. Several studies have suggested that inter-observer agreement increases with practice (Manske et al., 2002; Thomsen and Baadsgaard, 2006; Capion et al., 2008b). Annual standardization of the trimmers could be a way to unify records and keep quality at a high level. Until trimmers have standardized lesion identification and scoring, it will be incorrect to interpret trimming records without knowing how the individual trimmer is performing the records.

The trimmers were able to identify lesions at a satisfactory level based on photographs. However, they did not perform according to their ability under field conditions. There was a tendency that

lesions as HE, SH, and WL, which in mild cases seldom leads to lameness (Capion et al., 2009), had a lower recording rate than more obvious and severe lesions as SU, DD, and IH. This corresponds to the findings of Manske (2003). All three groups of trimmers showed this recording trend, and these findings may therefore apply for trimmers in Denmark and TTs in Ontario. The tendency was interpreted as trimmers having a lower motivation for recording clinically less important lesions under field conditions.

One thing is the ability to score lesions correctly, another is the motivation for doing it. Throughout the study, the author observed that trimmers with a high motivation for recording tended to make trimming records of higher quality compared with trimmers who had lower motivation, regardless of education. Trimmers who understood the reasoning behind making the records, and who had seen results of the effort appeared more motivated, which was also suggested by Eriksson (2006). Though the recording rate may be interpreted as the trimmer's attitude towards recording, it could also be a reflection of the farmer's attitude towards trimming records. If the latter is the case, the consistency of lesion scoring and recording is an issue that needs investigation. In herds where the farmer consider recording to be unnecessary, the record quality is probably lower compared to a farmer that encourages recording, as suggested by Bergsten (2010). To unravel what effects the attitude towards trimming records for both farmers and trimmers, more research is needed.

Education alone does not determine the quality of trimming records, and there seems to be more factors influencing the outcome. Higher quality of hoof health records will improve their value as a management tool for the herds, in breeding programs and future research.

Conclusion

The trimmers' lesion scoring abilities based on photographs was not reflected in the quality of their trimming records. There was no distinct difference in the quality of trimming records between the

three groups of hoof trimmers. However, there was a pattern of SU, DD, and IH being recorded at a higher rate compared to HE, SH, and WL irrespective of education. It appears that the attitude towards recording lesions determines the quality, irrespective of education and with no correlation to trimming experience, individual trimmer, time of data collection, or number of trimmers working together. Future studies are needed on factors affecting trimming record quality.

Conflict of interest statement

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

Acknowledgements

The author wishes to thank trimmers, farmers, friends, and family for their help with this thesis. A special thanks to my supervisor Nynne Capion for support and advice as well as Vic Daniel for exceptional hospitality, discussions, and encouragement. The study received funding from Viking Genetics, PLAN-Danmark, Oticon fonden, and Dansk Klovbeskærerforening.

References

Bergsten, C. 2010. Genetic achievements of claw health by breeding. WCDS Advances in Dairy Technology. 22:http://www.wcds.ca/proc/2010/Manuscripts/p349-356Bergsten.pdf.

Capion, N., A. M. Dansher and P. Nielsen. 2009. Klovsygdomme. Pages 6 in Klovsundhed Sund Klov i Malkekvægsbesætninger. Klovsygdomme. Dansk Landbrugsrådgivning Landscenteret, .

Capion, N., S. M. Thamsborg and C. Enevoldsen. 2008a. Conformation of hind legs and lameness in danish holstein heifers. Journal of Dairy Science. 91:2089.

Capion, N., S. M. Thamsborg and C. Enevoldsen. 2008b. Prevalence of foot lesions in danish holstein cows. Veterinary Record; 2008.163: 3, 80-86.

Daniel, V. 2012. Ontario dairy hoof health project.http://www.ontariohoofhealth.ca/OHHDfinal.pdf.

Eriksson, J. 2006. Swedish sire evaluation of hoof diseases based on hoof trimming records. Interbull Bulletin. 49.

Ettema, J., B. Nielsen and S. Østergaard. 2009. Økonomisk betydning af laminitis-relaterede sygdomme og digital dermatitis. Pages 28 in Klovsundhed Sund Klov i Malkekvægsbesætninger. Økonomisk betydning af laminitis-relaterede sygdomme og digital dermatitis. Dansk Landbrugsrådgivning Landscenteret Dansk Kvæg, Denmark.

Greenough, P. and et al. 2011. Lesion severity scoring guide.2013:http://www.ontariohoofhealth.ca/LesionSeverityGuide.pdf.

Kofler, J., R. Pesenhofer, G. Landl, I. Sommerfeld-Stur and C. Peham. 2013. Monitoring of dairy cow claw health status in 15 herds using the computerised documentation program claw manager and digital parameters. Tieraerztl. Prax. Ausg. Grosstiere Nutztiere. 41:31-44.

Landis, R. and G. Koch. 1977. The measurement of observer agreement for categorical data. Biometrics. 33:159-174.

Lowry, R. 2013. Kappa as a measure of concordance in categorical sorting. Http://Vassarstats. Net/Kappa. Html. 2013:.

Manske, T. 2003. Om klövvårdares klövhälsoregisteringar. Sven Vet Tidn. 15:11-18.

Manske, T. 2002. Hoof Lesions and Lameness in Swedish Dairy Cattle. Doctoral thesis Thesis, Swedish university of agricultural sciences, Skara.

Manske, T., J. Hultgren and C. Bergsten. 2002. Prevalence and interrelationships of hoof lesions and lameness in swedish dairy cows. Preventive Veterinary Medicine; 2002.54: 3, 247-263.

N. Capion. personal communication. DVM and Ph. D. Depertment of Large Animal Sciences, University of Copenhagen. Phone +45 35332843. E-mail nyc@life.ku.dk:.

NAV. 2011. Improved genetic evaluation -use of genomic information and claw health data. Press Release, Nordisk Avlsværdi Vurdering.

Nielsen, P. personal communication. DVM. Videnscenteret for landbrug, Kvæg. Phone: +45 87405286. Email: pin@vfl.dk:.

O'callaghan, K., P. Cripps, D. Downham and R. Murray. 2003. Subjective and objective assessment of pain and discomfort due to lameness in dairy cattle. Anim. Welfare. 12:605-610.

Seested, S. and M. Thomsen. 2011. Har du styr på dine klove -så får du råd til det sjove. Grobund Magasinet Om Liv Og Vækst På Landet. 2:4-http://www.e-pages.dk/videncentretforlandbrug/35/4.

Thomsen, P. and N. Baadsgaard. 2006. Intra- and inter-observer agreement of a protocol for clinical examination of dairy cows. Prev. Vet. Med. 75:133-139.

Van der Linde, C., G. de Jong, E. Koenen and H. Eding. 2010. Claw health index for dutch dairy cattle based on claw trimming and conformation data. J. Dairy Sci. 93:4883-4891.

Videnscenteret for landbrug. 2010. Nordisk klovatlas. Videnscenteret for Landbrug. https://www.landbrugsinfo.dk/Kvaeg/Sundhed-og-dyrevelfaerd/Klove-og-lemmer/Sider/Nordisk-Klovatlas.pdf?download=true:.