Swedish University of Agricultural Sciences
Department of Clinical Sciences

Association between animal-based measures and register-based welfare indicators in dairy cows – a study of the advisory service ‘Ask the Cow’ and the web report ‘Animal Welfare Signals’

Emilia Wallberg

Degree project, 30 credits
Advanced level, A2E
Agricultural Scientist Program - Animal Science Master Degree
Department of Clinical Sciences
Uppsala 2013

Swedish University of Animal Science
Faculty of Veterinary Medicine and Animal Science
Department of Clinical Sciences

Department of Clinical Sciences

Association between animal-based measures and register-based welfare indicators in dairy cows – a study of the advisory service ‘Ask the Cow’ and the web report ‘Animal Welfare Signals’

Samband mellan djurbaserade mått och välfärdsindikatorer i kodatabasen – en studie av rådgivningstjänsten ”Fråga Kon” och webbrapporten ”Signaler Djurvälfärd”

Emilia Wallberg

Supervisors:
Anki Roth, Swedish University of Agricultural Sciences, Uppsala
Louise Winblad von Walter, Växa Sverige, Stockholm
Marie Jansson Mörk, Växa Sverige, Stockholm

Examinator:
Ulf Emanuelson, Swedish University of Agricultural Sciences, Uppsala

Extent: 30 credits
Course title: Master Degree in Animal Science
Course code: EX0692
Program: Agricultural Scientist Program, Animal Science
Level: Advanced, A2E

Place of publication: Uppsala
Year of publication: 2013

On-line issue: http://epsilon.slu.se

Key words: dairy cows, animal-based measurements, register-based welfare indicators, resource-based measurements, ‘Ask the Cow’, ‘Animal Welfare Signals’, the Swedish Official Milk Recording Scheme, epidemiology, animal welfare, Växa Sverige, Swedish Dairy Association

Nyckelord: mjölkkor, djurbaserade välfärdsmått, välfärdsindikatorer i kodatabasen, resursbaserade välfärdsmått, Fråga Kon, Signaler Djurvälfärd, kokontrollen, epidemiologi, djurvälfärd, Växa Sverige, Svensk Mjölk

Preface

My interest in animal welfare has always been a strong motivation and guideline while studying to become an agronomist. I have read additional courses remotely to immerse myself in the subject and the aim of my studies has always been to work with animal welfare in some way. During my five years at the Swedish University of Agricultural Sciences (SLU) in Uppsala I have come in contact with the organisation of Växa Sverige (former Swedish Dairy Association) several times. I was always fascinated about their work; how they validate the importance of maintaining a good animal welfare and care a lot about the farmers at the same time.

In spring 2012 I contacted Louise Winblad von Walter and asked her if Växa Sverige needed any help with studying something associated to animal welfare. To my enthusiasm, she answered that they needed to investigate the association between the advisory service 'Ask the Cow' and the web report 'Animal Welfare Signals' further. Both were already applied as animal welfare tools, but it was necessary to convict their reliability. This was the beginning of a journey that evolved me as a person and enhanced my interest concerning welfare of dairy cows. This project has been educative and developing on many levels. Most of all it made me realize the complexity in evaluating animal welfare.

This project of 30 credits is part of a Master Degree in Animal Science and included in the Agricultural Scientist Program of Animal Science. Hopefully it will support Växa Sverige in their continuing work for improving welfare of dairy cows.

Pleasant reading!

Many thanks to:

Anki Roth; self-employed and researcher at SLU Uppsala, who as main supervisor has been supportive and helpful in every possible way during the whole process.

Louise Winblad von Walter; animal welfare at Växa Sverige, who as supervisor has been supportive and positive from the beginning until the end of the thesis.

Marie Mörk; animal health statistics at Växa Sverige, who as supervisor has been educational, patient and knowledgeable during all statistic analyses in particular but in all ways possible.

Växa Sverige (former Swedish Dairy Association) that made this project possible.

The Faculty of Veterinary Medicine and Animal Science at SLU, who saw the interest in being involved and made it possible for me to complete the project.

My friends and family who as always have been supportive and believed in me.

Table of Contents

Abstract

Animal welfare can be measured in several different ways; for example by using animal-based measurements. It has been shown that register data from the Swedish Official Milk Recording Scheme can be helpful when identifying herds at risk of having poor animal welfare, which led to a development of the advisory service of 'Ask the Cow' and the web report 'Animal Welfare Signals'. The aim of this project was to find out if register data from 'Animal Welfare Signals' can be used as a screening tool for detecting herds at risk of having poor animal welfare and in need of an 'Ask the Cow'. 'Ask the Cow' and 'Animal Welfare Signals' preferably should present the same view of reality; herds with most weaknesses in animal welfare should be identified in both systems. Welfare indicators in 'Animal Welfare Signals' were used as a test-kit for identifying herds at risk of having poor animal welfare according to the gold standard of 'Ask the Cow'. Herds were classified as having poor animal welfare if they had two or more welfare remarks and a herd got a welfare remark if they were among the 10% worst herds. Results indicate that limits for the 50^{th} and 90^{th} percentile currently applied within 'Ask the Cow' are not consistent with limits given by the study and therefore current limits need to be followed up. The highest found sensitivity was 0.73 and the highest found specificity was 0.71. All negative predictive values found in the final test-kits were considered high, and therefore herds with a good animal welfare are identified to a high degree. All of the final test-kits were not particularly consistent with each other, therefore it is recommended to apply several welfare indicators when identifying herds at risk of having poor animal welfare. However, the welfare indicator of *age at first calving* was found in all final test-kits and can therefore be considered particularly interesting. 'Ask the Cow' and 'Animal Welfare Signals' reflect the same view of animal welfare, but as result reviles it is recommended to use both of them when assessing animal welfare at herd level.

Sammanfattning

Djurvälfärd är ett komplext begrepp och kan bland annat definieras genom att använda djurbaserade mått på besättningsnivå. Det har tidigare bevisats att registerdata från den svenska officiella mjölkdatabasen Kokontrollen kan användas för att identifiera besättningar som riskerar att ha en bristande djurvälfärd. Detta ledde till utvecklingen av rådgivningstjänsten "Fråga Kon" samt webbrapporten ”Signaler Djurvälfärd”. Syftet med studien var att undersöka om registerdata från "Signaler Djurvälfärd" kan användas för att identifiera besättningar med brister gällande djurvälfärden och som är i behov av en "Fråga kon"-bedömning. Det är viktigt att se till att "Fråga Kon" och "Signaler Djurvälfärd" speglar samma bild av verkligheten, det vill säga det är önskvärt att besättningar med bristande djurvälfärd identifieras av båda systemen. Olika kombinationer av nyckeltal från "Signaler Djurvälfärd" användes för att identifiera besättningar med sämre djurvälfärd enligt guldstandarden ”Fråga Kon”. Till besättningar med brister gällande djurvälfärden räknades de som hade två eller fler avvikelser och en besättning fick en avvikelse om de var bland de 10 % sämsta besättningarna. Resultaten visar att de gränser som för närvarande tillämpas inom "Fråga Kon" för den 50:e och 90:e percentilen inte är förenliga med gränserna som studien tagit fram och de behöver därmed följas upp. Studien fann en högsta sensitivitet på 0,73 och en högsta specificitet på 0,71. Samtliga negativa prediktiva värden ansågs höga, och därmed identifieras en stor andel av de besättningar som har god djurvälfärd. Samtliga kombinationer av nyckeltal skiljde sig ifrån varandra, vilket innebär att det är viktigt att ett flertal nyckeltal tillämpas för att identifiera besättningar som riskerar att ha en bristande djurvälfärd. Nyckeltalet *inkalvningsålder* förekom dock i samtliga kombinationer av nyckeltal och kan därför antas vara av särskilt intresse. "Fråga Kon" och "Signaler Djurvälfärd" speglar samma bild av djurvälfärden, men utifrån resultaten i denna studie bör man använda båda systemen för att utvärdera djurvälfärden på besättningsnivå.

1. Introduction

1.1 Background

Animal welfare is important to consider in all industries where animals are used for any type of production. This can be motivated in many ways, for example by claiming that it is a human responsibility to ensure good animal welfare when using animals for human sake (Mellor *et al*, 2009). Another aspect is that consumers become more and more aware about the concept of animal welfare; they demand more of the animal industry (Blokhuis *et al*, 2003). Farmers can have a more profitable production as a direct consequence of improved animal welfare (Hulsen, 2011).

There are different methods for measuring animal welfare and animal-based measures is one of them (Welfare Quality®, 2009). In a study by Sandgren *et al* (2009), it was suggested that register data from the Swedish Official Milk Recording Scheme (SOMRS) could be helpful to identify herds at risk of having poor animal welfare. Results from animal-based measurements were compared with associated case herd data from SOMRS. The study found a final set of welfare indicators based on data from "SOMRS" that correctly could classify 77% of the herds at risk of having poor animal welfare.

Based on results from the study by Sandgren *et al* (2009), two animal welfare services were developed; 'Animal Welfare Signals' and 'Ask the Cow'. 'Animals Welfare Signals' is a web report consisting of 24 register based welfare indicators associated with animal welfare or economy; available for all herds participating in SOMRS (Svensk Mjölk, 2011[1] and Mörk, 2001). Results for each herd are presented as different coloured smileys (Appendix 2). 'Ask the Cow' is an advisory tool where specially trained assessors use animal-based measurements to objectively evaluate animal welfare on farm (Mörk 2011). 'Ask the Cow' identifies welfare issues in need of improvements and can be applied on all types of herds. Results are presented as flowers with petals of different lengths (Appendix 3) (Thunberg and Winblad von Walter, 2011).

It is important to ensure that 'Ask the Cow' and 'Animal Welfare Signals' reflect the same view of reality. This means that herds with most weaknesses in animal welfare should be identified in both systems. If not, there is risks of herds with impaired animal welfare remain unnoticed or that farms with good animal welfare become falsely identified as farms at risk of having poor animal welfare. Therefore it is necessary to evaluate how well both systems are consistent with each other and if any adjustments or changes are needed in order to ensure their credibility.

1.2 Aim

The aim of this thesis was study if the advisory service 'Ask the Cow' and the web report 'Animal Welfare Signals' reflect the same view of reality; namely animal welfare in Swedish dairy cows. Thereby, this project could convict if the web report is reliable for detecting weaknesses of animal welfare within Swedish dairy herds.

Following questions would be answered during the project:

1) Do ‘Ask the Cow’ and ‘Animal Welfare Signals’ reflect the same view of animal welfare?

2) Can ‘Animal Welfare Signals’ be trusted as reliable assessments for detecting strengths and weaknesses of animal welfare?

3) Is it possible to distinguish a set of welfare indicators in ‘Animal Welfare Signals’ that are additional important?

1.2.1 Hypothesis

Herds regarded as having poor welfare by ‘Ask the Cow’ can be identified by a set of register based welfare indicators in ‘Animal Welfare Signals’. As a consequence of that, ‘Animal Welfare Signals’ can be used as a screening tool to detect herds with risk of having poor animal welfare.

1.3 Limitations

Herds included in the project were not randomly selected. All herds in which an ‘Ask the Cow’ had been performed were included and the reasons for them performing an ‘Ask the Cow’ was unknown.

2. Literature

2.1 Animal Welfare

Animal welfare is a complex and challenging concept that can be defined with somewhat different declarations. Fraser (1989) stated that there are five main factors related to health, husbandry and ethical concerns;

- "e*thical use of the animal,*
- *standards of husbandry and production which meet an attainable level,*
- *control of suffering for the wellbeing of the animal,*
- *provision of veterinary care,*
- *and ecological management"*

Animal welfare can also be defined as a combination of three combined perspectives; animals' natural way of living, animals' biological performance and animals' feelings (Fraser *et al,* 1997). For example, the perspective of animals' natural way of living can be declared by their natural behaviours, which of course differs between species and animals. The perspective of animals' biological performance can be declared by words like fitness and stress (Broom, 1998). The perspective of animals' feelings refers to animals' emotions and what they feel (Fraser *et al,* 1997).

Another way of defining animal welfare is in terms of "the five freedoms", adapted by the Farm Animal Welfare Council (FAWC, 2009; Fraser, 2008). These are;

- "*freedom from hunger and thirst,*
- *freedom from discomfort,*
- *freedom from pain, injury or disease,*
- *freedom to express normal behaviour*
- and *freedom from fear and distress".*

Furthermore, if solely ethical concerns that affect animals' quality of life are to be discussed Fraser *et al* (1997) suggested that animal welfare could be concerned by three different sorts of problems. Problem number one are "*Adaptions that no longer serve an important function*" (e.g. calves drinking milk from a bucket but are strongly motivated to suckle), number two are "*Challenges for which the animal lacks corresponding adaptions*" (e.g. animals who breath contaminated air may not directly notice that, but it will lead a disease later in life) and number three are "*Challenges for which the animal has corresponding adaptions*" (e.g. animals can adapt with temperature, but at a certain point they will be negatively affected by it). Depending on which one of these factors, explanations and/or problems one choose to follow the concept of animal welfare can have different meanings for different people (Fraser, 1989).

Animal welfare is not mainly motivated by economy or scientific interests but is also driven by social and ethical concerns (Rushen *et al,* 2008). Since concerns about animal welfare differ among stakeholders (consumers, general public, policymakers), efforts to improve

animal welfare must meet all expectations. In general, consumers' interest for the process behind food products and the wellbeing of production animals has increased (Blokhuis *et al*, 2003). One specific reason for that are probably disease outbreaks (foot-and-mouth disease, swine fever etcetera), but also an increased interest for food safety and the welfare of animals. Interest in farm animal welfare usually increases as countries develop (Rushen *et al*, 2008). As countries become more financially secure, food industry develop and intensive production of farm animals increases. With an intensive farm animal system, consumers are more or less required to consider animal welfare issues. Consumers' consciences were challenged as intensive production systems arose out of the desire to be effective and profitable. As an example, intensive farming is often associated with animals kept indoors with limited possibilities for expressing their natural behaviours. This may affect the general public and consumers that start to think more about how their food has been produced.

2.2 Animal Welfare in Cattle

Within cattle production certain animal welfare concerns are often discussed. For example, this applies for early cow-calf separation, rough handling, transport, dehorning and absence of grazing (Rushen *et al*, 2008). There is a clear connection between health and animal welfare. As a consequence of that, diseases or physical damages are often used as a way of assessing animal welfare in cattle. However, Rushen *et al* (2008) also emphasizes the fact that it can be difficult to determine to what extent an animal that is injured or ill really suffers from poor animal welfare. On the other hand, it has been proven that cows produces less milk and decrease their feed intake as a consequence of different diseases (Bareille *et al*, 2003). It is also important to stress that a high production not necessary is evidence for a good animal welfare (Rushen *et al*, 2008). As an example, if growth enhancers are used in beef cattle production they may grow regardless of whether animal welfare is poor or not. Depending on type of flooring, an animal welfare issue such as lameness can be more or less common (Albright and Arave, 1997). For example, appearance of lameness increases if the floor is unclean or made of slippery concrete.

Another aspect of animal welfare in cattle is stress, which can be assessed using physiological measurements (Rushen *et al*, 2008). However, there are several constraints that need to be taken into account if using physiological measurements. For example, when prolonged stress is to be measured it can be difficult to determine associated changes in the physiological system. An additional way of assessing animal welfare is by studying animals' behaviour. For example there are methods for measuring preferences of the animal, but signalling behaviour, abnormal behaviour and playing behaviour are also of interest when animal welfare is assessed. Animals' natural behaviour can serve as a guidance to prevent specific animal welfare problems. According to Rushen *et al* (2008), animal welfare is compromised if the animals cannot express their natural behaviours.

It is important to observe the animals and interpret the signals correctly (Hulsen, 2011). In the Cow Signals® concept, three questions are asked during observations; "*What do I see? Why do I see it? What does it mean?*". Observing is to be genuinely aware about what the animal is signalling and apply this knowledge for making improvements. It is important to observe

both the herd as a whole unit as well as individual animals; each measurement will communicate specific signals. There is always a reason for a cow's behaviour and it is important to consider current circumstances in order to evaluate the cause. In order to evaluate behaviours it is important to know how it is supposed to look like; meaning what is normal. If a farmer want to improve the production it is vital to find faults and avoid risks. In Sweden, trained assessors educate for instance farmers and advisors in Cow signals® with the aim of improving animal welfare, production and economy on herd level (Roth, personal communication, 2013).

2.3 Legal Aspects

In order to express legal aspects of animal welfare it is necessary to look at it in a scientific point of view (Broom, 1991). All member states of the European Union must follow the "Council Directive concerning the protection of animals kept for farming purposes" (98/58/EC). This directive is designed for protection of farm animals. The "European Convention for the protection of animals kept for farming purposes" is more specific than the directive and is specialized on farm animals kept in intensive production systems (L323, 1978). Furthermore, there are directives and regulations that specifies in individual species and/or on particular handling activities within the production chain. Within cattle production a dedicated directive protects calves according to the "Council Directive of minimum standards for the protection of calves" (2008/119/EC). In present time there are no other directives that specifically protect other animal categories within the cattle production. Both the directive and convention are minimum standards and each member state can decide to apply stricter rules within the country. Compared to many other European countries, Sweden applies stricter rules concerning animal welfare. This is primarily governed by "The animal welfare act" (SFS 1988:534) as well as "The animal protection ordinance" (SFS 1988:539). Additionally, farm animals in Sweden are protected by regulations and guidelines stated by the Swedish Board of Agriculture (SJVFS 2010:15).

One challenging part of maintaining a high animal welfare standard is to assess it. The Welfare Quality® project was funded by the European Union (Welfare Quality®1 and Welfare Quality®2, 2004). The project started in 2004 and finished about five years later. In order to really measure the animal welfare status, on-farm monitoring systems are needed (Blokhuis *et al*, 2003). Potential risks that could affect animal welfare are evaluated in this type of monitoring system. For example, transports and slaughterhouses are included in the concept of "on-farm" monitoring and therefore potential welfare risks could be prevented in all stages of the production chain. One aim of the Welfare Quality® project was to improve the quality of life for farm animals in Europe and to develop the European standard of minimal animal welfare; it should be possible to "*assess and monitor animal welfare in production*" (Welfare Quality®1 and Welfare Quality®2, 2004). One substantial part of the Welfare Quality® project was to integrate animal welfare in the whole food quality chain. It was also important to make sure that both herd managers and the market could take advantage of the results. The project mainly concentrated on three species; cattle, pigs and chickens, as well as related animal products. After the Welfare Quality® project was completed, each specie had a corresponding assessment protocol.

The Welfare Quality® assessment protocol for cattle mainly practices animal-based measures as assessment tools (Welfare Quality®, 2009). The protocol principally contains measures for both dairy cows and beef cattle. Animal-based measures focus on the animal itself, meaning animals function as a model for all measures. For example, this means that the assessor evaluate lameness and body condition in cows. However, resources-based and management-based measures are not excluded from the assessment protocol. Resource-based measures focus on the environment of the animal, the environment then function as a model for all measures. For example, this means that the assessor evaluate presence of straw and box sizes. Management-based measures focus on manager processes and how the manager handles the animal unit, which for instance means that the assessor evaluate what the herd manager does to prevent animal diseases. The Welfare Quality® project recognized four welfare principles; good feeling, good housing, good health and appropriate behaviour (Forkman and Keeling, 2009) (Appendix 1). The principles were converted into twelve welfare criterias, which in turn were linked to several on-farm measurements (Forkman and Keeling, 2009; Winckler et al, 2007). Within the monitoring system of dairy cows, the measurements are adapted to the specific specie and production system. For example dairy cow measurements contain body condition and cleanliness of udder (Welfare Quality®, 2009).

According to a project by Knierim and Winckler (2009), where parts of the Welfare Quality® project is evaluated, animal-based measures of animal welfare meet some misfortunes regarding its feasibility. The animal-based measurement can be time consuming. However, some of the Welfare Quality® projects innovative solutions, such as the use of administered databases, also may save time and in that way improve animal-based measures. Additionally, the Welfare Quality® project convict that the probability for validity often are considered high in animal-based welfare measures as they are measured by experts (Scott *et al*, 2001). However, Knierim and Winckler (2009) also stress the importance of always trying to improve the assessment system through for example assessors training. It is important to clarify that the Welfare Quality® assessment protocols only function as a supplementary tool for animal welfare assessment and that each country has its own animal welfare rules to follow (Welfare Quality®, 2009).

2.4 Official Milk Recording in Sweden

The Swedish Official Milk Recording Scheme is a national database for dairy cows. It includes information from approximately 85% of all Swedish dairy cows (Svensk Mjölk, 2011[1]). It is a reliable system where the producer can register and find essential information on his or her individual cows. The database is useful for the producers as well as its caretaker Växa Sverige (former Swedish Dairy Association) and other organizations within the dairy sector (Svensk Mjölk, 2012[1]).

Selected welfare indicators from SOMRS have been proven to be associated with animal-based measures in the herd when monitoring animal welfare (Sandgren *et al*, 2009). The study indicates that it might be economically advantageous to initially, before using animal-based measures, identify herds with risk of having poor animal welfare according to figures from SOMRS. In the study by Sandgren *et al*, nine animal-based measures were selected and

examined in 55 Swedish dairy herds (18 kept in loose housed systems and 37 in tied stalls). Cleanliness and body condition were examined in all animal categories; calves, young stock and dairy cows. The animal-based measures of lameness, injuries and inflammations, rising behaviour and avoidance distance were only examined in cows. Herds were classified as having poor animal welfare if they were among the 10% worst herds in two or more of the included animal-based measures, and otherwise as not having poor animal welfare. Thirteen herds were counted as having poor animal welfare (23.6%) and represented "the gold standard". Several animal health experts' selected 65 key figures from SOMRS as possible indicators of analysing animal welfare. Different combinations of key figures were then tested against the gold standard.

Finally the study could point out a set of welfare indicators that properly could categorize 77% of the herds at risk of having poor animal welfare (Sandgren *et al*, 2009). The figures were calf, young stock and cow *mortality* as well as *heifers not bred >17 months* and *cows with late on-going artificial insemination (Als), >120 d* (or "*calving - latest insemination, >120 d"*). Result showed that it is possible to use pre-collected register data, such as SOMRS, when herds at risk of having poor animal welfare are to be found. Sandgren *et al* claims that similar types of pre-collected data advantageously can be used in other countries as well. However, each country needs to adjust the method in accordance to local conditions, which means that selected welfare indicators must be adapted to local on-farm animal-based measures. Based on the study by Sandgren *et al*, the Swedish Dairy Association developed an animal welfare advisory tool and an animal welfare web report; 'Ask the Cow' and 'Animal Welfare Signals' respectively (Mörk, 2011). Växa Sverige is now responsible of both these systems.

2.5 'Animal Welfare Signals'

'Animal Welfare Signals' is based on animal welfare related welfare indicators from SOMRS, (Mörk, 2011) (Appendix 2). It is a web report aimed at presenting animal welfare status at herd level in a comprehensive way and the farmer has access to his or her web report. In a benchmarking system, it demonstrates the animal welfare status of the specific herd in relation to other herds in the system. 'Animal Welfare Signals' includes 24 welfare indicators divided into seven different focus areas; calves, heifers, calving, feed balance, diseases, monitoring and management as well as sustainability (Appendix 2) (Svensk Mjölk, 2011[2]). These welfare indicators are for example *calf mortality, percentage of cows with low urea values* and *calving interval*. Behind each welfare indicator there are formulas and codes processing numbers from each input data; data from individual animals in each herd (Svensk Mjölk, 2011[3]). Numbers are partly documented by the farmer, but most of them are reported automatically into the system.

Each welfare indicator is presented with a smiley for that welfare indicator, a yellow smiley implies that the herd are among the 10-50% worst herds and herds among the 10% worst are assigned with a red smiley (Appendix 2) (Mörk, 2011). Herds among the 50% best are assigned with a green smiley instead. The herd figures for all welfare indicators for the latest three and twelve months and the latest milk-recording year are presented in a table (Svensk

Mjölk, 2011[2]). Also, to enable comparisons with other herds, the figures for the 10% best, the median and the 10% worst herds in the SOMRS are presented in the table. The twelve-months figure, which represents production data from a whole year back in time, determines the colour of the smiley. The three-months figures are based on data from the past three months and reasoning behind the shorter time interval is to highlight this period close in time to discover possible negative trends. When presented in 'Animal Welfare Signals', those numbers based on facts from the past three months is presented as a future-trend after having undergone specific calculations.

2.6 'Ask the Cow'

'Ask the Cow' is an advisory service where animal-based welfare measures such as *body condition, cleanliness* and *lameness* etcetera are used to objectively assess animal welfare at herd level (Appendix 3) (Mörk, 2011). According to currently applied assessment protocol, the assessment point of *rumen fill* is voluntary for the assessor to measure (Svensk Mjölk, n.y.[1]). A calibrated assessor performs the assessment and the result is presented in an overall picture of animal welfare at herd level (Mörk, 2011). To become a calibrated assessor, it is necessary to participate in a course and yearly updates (Thunberg and Winblad von Walter, 2011). It is important to emphasize that results given by 'Ask the Cow' reflects animal welfare in a specific herd at a particular date (Winblad von Walter, personal communication, 2012). The assessment takes about two to three hours to perform and includes calves, young stock as well as cows (Svensk Mjölk, n.y.[2]). The result of an assessment is compared to the results of the other performed assessments and presented as a table and two pictures of flowers; one flower for calves/young stock and one for cows (Appendix 3). If the result indicates poor animal welfare, petals of the flowers are broken. If a petal is half its length the herd is among 10-50% worst herds. If it is even shorter the herd is among the 10% worst herds (Thunberg and Winblad von Walter, 2011). For example, the limit for the 90th percentile is 15% for the assessment point of *injuries* in cows and this means that herds with 15% or more injuries are qualified for this group (Winblad von Walter, personal communication, 2013). After the assessor is done with analysing the results, the farmer is contacted and provided with an action-plan with practical advises of how to improve animal welfare at herd level (Svensk Mjölk, n.y.[2]). The suggestions should help the farmer to improve not only animal welfare but herd profitability as well.

'Ask the Cow' can be implemented in a herd after a request from for example a veterinarian, an advisor or the farmer (Thunberg and Winblad von Walter, 2011). Also, certification companies such as for example Swedish "KRAV" can request it as part of improving animal welfare at herd level (KRAV, 2013). Because of its animal-oriented focus "Ask the Cow" is suitable for all types of production systems; cows in loose housed systems or tied stalls as well as old or newly built barns (Thunberg and Winblad von Walter, 2011). 'Ask the Cow' should be performed one or two hours before milking and after feeding because this is considered as the most peaceful time period of the day. During the assessment, 35 animals of each group (calves, young stock and cows) are randomly selected and estimated. However, the number can sometimes be smaller depending on herd size. First, the whole group is registered and executive assessor notes presence of for example vermin, ringworm, social

trimming as well as number of water cups and feed places etcetera. Thereafter 35 individual animals are assessed within each animal category. If the group consists of more than 35 animals they are randomly selected. To begin with 35 cows are assessed for *body condition, rising behaviour, cleanliness, coat, claws, lameness, injuries, rumen fill* and *injury placement*. For example, the assessment point of *rumen fill* signals if the cow has been eating and how well the feed degradation has been functioning during the last couple of hours or even days (Hulsen, 2011). The assessment point of *rising behaviour* signals how well the lying area is adapted to the size of the cow and how comfortable the cow is with the space available. The cow assessment is followed by assessments at group level of young stock and calves, where for example presence of cough, diarrhoea and deviant behaviour are examined (Thunberg and Winblad von Walter, 2011). Thereafter 35 individuals of young stock and calves each are registered for *body condition, cleanliness, injuries* and *health*. Before leaving the herd executive assessor usually review the results from 'Ask the Cow' verbally for the farmer.

Usually executive assessor follows up the assessment with a revisit (Thunberg, and Winblad von Walter, 2011). He or she returns to the herd to discuss specific animals that have received a deviation and/or possible arrangements for improving animal welfare in general. This type of revisit is supposed to provide the producer with tools for detecting animal welfare issues in time. During revisit the book "Cow Signals" is usually applied (Hulsen, 2011; Thunberg, and Winblad von Walter, 2011). It is a practical book for farmers within the dairy cow industry and as the title indicates it contains the Cow Signals® concept (Hulsen, 2011). It focuses on the animal itself and what type of signals the animal communicates. Furthermore, as parts of the animal welfare process an additional 'Ask the Cow'-assessment is preferred (Thunberg, and Winblad von Walter, 2011).

2.7 Diagnostic Tests

Epidemiology is a method for explaining the cause, range and progression of diseases within a population (Thrusfield, 2005). Within epidemiology it is possible to identify to what extent and how a disease is part of a population. Within epidemiology diagnostic tests are practiced and these tests are a method for distinguishing between sick and healthy individuals. The result from a diagnostic test can be divided into two groups; positive and negative test results (Appendix 4). The true positive test result could represent that a sick cow is positively tested for the specific disease, and the true negative test result represent that a healthy cow are negatively tested for the same disease. It is important to clarify a precise cut-off point; exactly when the cow is counted for as sick/positive or healthy/negative in a continuous variable.

In order to evaluate a diagnostic test it is necessary to analyse its sensitivity and specificity (Thrusfield, 2005). Sensitivity is the ratio of sick individuals that also test positive and specificity is the ratio of healthy individuals that also test negative (Appendix 4). Sensitivity and specificity are a proportion and can therefore be stated as a range between 0-1 or in percentage. In order to avoid false positive and false negative values it is desirable to have as high sensitivity and specificity as possible. One precondition for calculating sensitivity and specificity is the presence of a truth to compare the test with; a "gold standard" (Thrusfield, 2005). Within the gold standard there are only a number of true positive and true negative

results (Appendix 4). However, within the test or “test-kit” there will be a number of true and false positive as well as true and false negative results.

The positive and negative predictive value of a diagnostic test is also of interest within epidemiology (Thrusfield, 2005). It is a way of measuring how well a test can predict if an animal is sick/positive or healthy/negative. A positive predictive value (PPV) is the ratio of tested individuals that are tested positive by the test and are truly sick, and negative predictive value (NPV) is the ratio of tested individuals are tested negatively by the test and truly are healthy (Appendix 4). Predictive values are a proportion and therefore expressed in values of 0-1. They are influenced by sensitivity, specificity and prevalence; if prevalence decreases, so does the positive predictive value. If prevalence increases, so does the negative predictive value. Sensitivity and specificity states the pulse of the test, while a predictive value states the pulse of participating individuals. When tests are applied on herd level, multiple testing is preferable. As the name implies, multiple testing consists of at least two tests. One of these methods is called parallel testing were sensitivity will increase and specificity will decrease. This will also lead to an increased negative predictive value. During parallel testing, the animal is tested in several tests at once and only needs to be positively tested in one of them to be affected. When parallel interpretation is applied there will be more false positive analysis than the opposite; that is a low PPV. Another example of multiple testing is serial testing, were one test at a time is conducted. Within serial testing only animals with a positive test result are tested once more. With this method, sensitivity and negative predictive value will decrease.

3. Material and Method

3.1 Data

Data was obtained from the advisory tool 'Ask the Cow' and the web report 'Animal Welfare Signals' at Växa Sverige in November 2012. Initially all performed 'Ask the Cow' assessments done since the service started in November 2009 until September 2012 was included (147 assessments in 135 herds). Assessments performed during education of 'Ask the Cow' advisors were not included. In one herd, two assessments had been performed at same date and one was therefore deleted (by giving each of them a random number and delete the assessment with the highest number). Data from 'Animal Welfare Signals' for three and twelve month prior to the assessment was then retrieved for all herds with an 'Ask the Cow' assessment. Three months data from the 'Animal Welfare Signals' was available for 129 of the assessments (119 individual herds) and twelve months data was available for 126 of the assessments (116 individual herds). The reason for missing data in 'Animal Welfare Signals' was that some herds were not enrolled in, or had recently joined, SOMRS. Three herds in 'Animal Welfare Signals' did not have data available for twelve months, and therefore the total number of herds differed. The analyses were performed in two different time intervals; three and twelve months. Moreover, both limits currently applied in 'Ask the Cow' (to point out welfare problems, currently used as limits for flowers with petals of half its length and "stumps") and limits given by the data material were used. All data editing and statistical analysis were conducted in the statistical software program of Stata® version 12 (Stata Corporation, College station, TX, USA).

3.2 Assessment Points from 'Ask the Cow'

3.2.1 Data editing

Assessment points and rating scales used in 'Ask the Cow' are included in Appendix 3. For cows, the following assessment points were used: *rising behaviour, body condition, cleanliness, coat, claws, injuries, lameness* and *rumen fill*. For young stock and calves the corresponding assessment points were *body condition, cleanliness, injuries* and *health*. The assessment point of *location of injury* was excluded for all of the animal categories due to practical reasons; since the material already was extensive. The assessment point of *claws* included both *claw length* and *outwardly angled claws*, which were treated as two separate assessment points. Similarly, *coats* included both *dusty neck* and *scuffed neck*, which was also treated as two separate assessment points. Available data from 'Ask the Cow' were the number of animals per rating for each assessment point and animal category, i.e. for example 5, 20 and 5 cows with rating lean, normal and fat for the assessment point *body condition*. The total number of assessed animals per assessment point and herd was calculated by adding up the number of animals per rating. The total number of estimated animals should not exceed 35, however some assessment points deviated from that number. This was explained by counting an animal more than once in the assessment points of *coats* and *claws*. For example, a cow can have long and pointy claws at the same time and thus be counted for twice within the assessment point of *claws*. If the deviant total number seemed reasonable (e.g. 34 cows without claw remarks plus one cow with long claws and one with pointy claws) it was

assumed that 35 cows had been assessed. As a consequence of that, the total number of assessed animals was changed to 35 for the assessment points *coats* and *claws* in one and 18 herds respectively. Nine herds had unreasonable total number of assessed animals for at least one assessment point and were excluded from further analysis.

Thereafter, the prevalence for each assessment point was calculated as total number of deviations divided with total number of estimated animals. The assessment points of *body condition, cleanliness* and *injuries* were divided into *lean* and *fat*, *dirty* and *severely dirty*, and *mild injuries* and *severe injuries*, respectively, for all animal categories. Prevalences were also estimated for the divided assessment points.

3.2.2. Identifying Herds With Poor Animal Welfare According to 'Ask the Cow'

Two definitions for welfare remarks were used, 1) based on the distribution of prevalence's in the data and 2) based on the limits currently applied in 'Ask the Cow'. The definitions of welfare remarks were used separately;

1) For each assessment point a herd was given a welfare remark if it was among the herds with 10% highest prevalence for that remark.
2) For each assessment point a herd was given a welfare remark if the prevalence exceeded the limit for being among the 10% worst according to currently limits used in 'Ask the Cow' (a stump in the welfare flower).

Specific assessment points were selected from 'Ask the Cow' and used as a gold standard for animal welfare; *lean*, *severely dirty* and *severe injuries* for all animal categories as well as *mild* and *severe lameness* for cows. Since some herds had missing prevalence within the gold standard of 'Ask the Cow', the total numbers of assessed herds were 113 in three months and 112 in twelve months. A herd was defined as having poor animal welfare if they had two or more welfare remarks in 'Ask the Cow'. Project limits and current applied limits in 'Ask the Cow' were used separately, as well as data from 'Animal Welfare Signals' in three and twelve months. The assessment points of *rising behaviour* and *rumen fill* in cows had few assessed herds and *health* in young stock and calves had low prevalence and they were therefore not included in the analysis.

3.3 Welfare Indicators from 'Animal Welfare Signals'

Data from 'Animal Welfare Signals' were retrieved for the periods three and twelve months before an 'Ask the Cow' was performed. Welfare indicators included in 'Animal Welfare Signals' are presented in Appendix 2. Initially data that extended over a period of twelve months in 'Animal Welfare Signals' were analysed, and the same statistical procedure followed for three months data. The time period of twelve months implies that data has been collected from one year before 'Ask the Cow' was performed, and the time period of three months implies that data has been collected from three months before 'Ask the Cow' was performed in that specific herd. The time period is therefore specific for every herd included. For example if 'Ask the Cow' was performed in September in one herd, the time period of

three months equals to data from ‘Animal Welfare Signals’ in June, July and August for that specific herd. Herds among the 10% worst herds within a welfare indicator got a welfare remark for that specific indicator.

3.4 Correlations

Spearman´s rank correlation was used to calculate correlations between the assessment points in ‘Ask the Cow’ and welfare indicators in ‘Animal Welfare Signals’ in three and twelve months respectively (Appendix 2 and 3)**.** This was done in order to see if any strong and statistically significant correlations could be found. For statistically significant correlations ($p \leq 0.05$), a scattered graph was created in order to explore the correlation. Additionally, every welfare indicator was analysed for possible equal values regarding number of welfare remarks in three and twelve months.

3.5 Use of Welfare Indicators to Detect Herds with Poor Welfare

The welfare indicators in ‘Animal Welfare Signals’ were used as a test-kit for identifying herds at risk of having poor animal welfare according to the gold standard ‘Ask the Cow’. The test-kit was similar to parallel interpretation of tests, i.e. a herd would be counted as having poor animal welfare if it had a remark in any of the indicators included (Thrusfield, 2005). First each welfare indicator was tested separately and the sensitivity and specificity for each welfare indicator were calculated. Thereafter the ambition was to found a kit of welfare indicators from ‘Animal Welfare Signals’ that when tested against the gold standard created a high sensitivity but still retained a relatively high specificity. The test-set was constructed by selecting welfare indicators with the highest sensitivity. After initial sensitivities and specificities were preformed, the welfare indicator with highest sensitivity was selected and combined with all of the other welfare indicators separately. Thereafter another welfare indicator that as a pair created the highest sensitivity and specificity was added. These two were again tested against each of the other welfare indicators and an additional welfare indicator with the highest numbers of sensitivity and specificity was chosen, and so on. The test evaluation was stopped at a combination of five welfare indicators. Even though the sensitivity could still improve, adding more indicators would only lower the specificity (Mörk, personal communication, 2013).

The final test-kit with a combination of five welfare indicators was then further tested. Analyse were completed by changing each limit for the included welfare remarks in the test-kit of welfare indicators to the 80^{th}, 85^{th} or 95^{th} percentiles instead. The purpose was to found out if this could improve the value of sensitivity and specificity. This was done separately in one welfare indicator at a time; first one limit was changed in one welfare indicator three times (80, 85 and 95%) at the same time as the other assessment points included in the test-kit were kept at the limit of the 90^{th} percentile. This procedure was accomplished for all five of the included welfare indicators and finally a combination of the highest sensitivity and specificity was selected.

4. Results

4.1 Descriptive Data

A total of 129 herds had data in both ‘Ask the Cow’ and ‘Animal Welfare Signals’ (126 herds for twelve month data). Of those 129 assessed herds, 90 herds had loose housed systems and 39 had tied-up stalls. Included ‘Ask the Cow’-assessments had been performed by about 20 different assessors with a range of 1-20 assessment each. All assessment points were not assessed in all herds; the highest number of assessed herds was 128 and the lowest 16. In Table 1 the total number of herds within each specific assessment point is shown. Table 1 also presents the 10^{th}, 50^{th} and 90^{th} percentile for each assessment point as well as current limits for flowers with a petal half of its lengths and a “stump” (equivalent to the 50^{th} and 90^{th} percentiles, respectively). Figure 1 presents the number of herds per number of welfare remarks for herds with data for all assessment points but *rumen fill*, and *rising behaviour* in cows and *health* in young stock and calves.

In Table 2, the 10^{th}, 50^{th} and 90^{th} percentile for all welfare indicators in ‘Animal Welfare Signals’ in twelve and three months respectively are presented. Figure 2 present numbers of herds per number of welfare remarks for herds in twelve and three months respectively. Number of herds differs due to that all herds did not have data for twelve months.

Table 1. Total number of evaluated herds within each assessment point, and the 10th, 50th and 90th percentile of the prevalence (%) in the study herds (project limits), and the 50th and 90th percentile currently used as limits (%) for flowers with petals of half its length and "stumps" in 'Ask the Cow' (current limits).

		Project limits			Current limits	
Assessment points	**Herds**	**10%**	**50%**	**90%**	**50%**	**90%**
Cows						
Rising behaviour*	95	-	-	-	-	-
Body condition	128	0	6	19	-	-
Lean	128	0	3	9	3	15
Fat	128	0	3	13	10	20
Dirty	128	0	11	46	25	60
Severely dirty	128	0	0	8	2	10
Dusty neck	128	0	6	63	-	-
Scuffed neck	128	0	0	0	-	-
Claw length	123	0	0	22	20	40
Outwardly angled claws	123	0	15	43	30	41
Lameness	128	0	6	23	5	15
Mild injuries	128	6	23	57	6	15
Severe injuries	128	0	3	17	3	7
Rumen fill*	16	-	-	-	-	-
Young stock						
Body condition	118	0	3	16	-	-
Lean	118	0	0	9	11	30
Fat	118	0	0	9	-	-
Cleanliness/dirty	117	0	14	52	36	75
Severely dirty	117	0	0	9	5	20
Injuries	118	0	0	10	0	6
Severe injuries	118	0	0	3	0	5
Health**	116	-	-	-	-	-
Calves						
Body condition	121	0	0	23	-	-
Lean	121	0	0	20	20	40
Fat	121	0	0	0	-	-
Cleanliness/dirty	121	0	9	40	24	47
Severely dirty	121	0	0	5	3	10
Injuries	121	0	0	5	0	6
Severe injuries	121	0	0	3	0	5
Health**	119	-	-	-	-	-

*Not included for further analysis due to the low number of herds in which those were assessed

**Not included for further analysis due to practical reasons

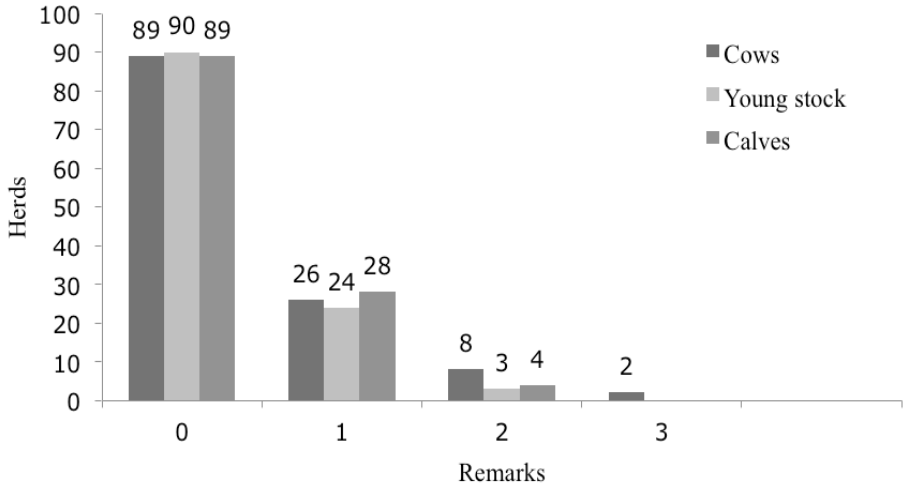


Figure 1. Number of assessed herds per number of welfare remarks (i.e. herds being among the 10% with the highest prevalence) and animal category. Totally there were 125, 117 and 121 herds with data on all assessment points for cows, young stock and calves, respectively

Table 2. Welfare indicators from 'Animal Welfare Signals' with the 10th, 50th and 90th percentile, for 126 herds with data for twelve months and 129 herds with data for three months.

	Twelve months			**Three months**		
	Percentiles			**Percentiles**		
Welfare indicators	10%	50%	90%	10%	50%	90%
Calf mortality, 1-60 days *	0	1.4	5.8	0	0	10.0
Calf mortality, 2-6 months *	0	0	4.5	0	0	4.9
Young stock mortality, 6-15 months *	0	0	4.5	0	0	6.3
Culling/mortality, first parity, 1-90 d after calving *	0	3.4	8.8	0	0	4.3
On-farm mortality **	1.6	5.1	10.4	0	4.9	13.0
Cows with diseases, total **	7.9	29.1	59.6	2.9	26.7	62.1
Low urea values (%)	1.7	5.7	17.7	0.6	5.5	20.8
Abnormal urea values (%)	12	18.7	29.0	8.8	16.6	32.8
Calving interval (months)	12.3	13.1	14.6	12.2	13.1	15.2
Incidence of calving difficulties **	1.2	5.0	11.8	0	4.7	14.3
Stillbirth, 0-24 hours *	2.1	6.4	10.9	0	6.0	14.3
Heifers not bred, > 17 months of age *	3.4	20.9	57.7	2.8	23.0	71.0
Culling total **	25.7	35.2	45.7	12.4	33.4	62.2
Incidence of Incidence of other feed-related diseases **	0	1.2	5.6	0	0	6.2
Culling, udder diseases **	2.7	9.6	18.2	0	8.5	26.5
Incidence of mastitis treatments **	1.8	14.1	28.3	0	11.2	36.1
Somatic cell counts from test milking (1000 cells/ml)**	145.0	217.0	328.0	135.0	225.0	353.0
Incidence of claw and leg disorders **	0	1.7	11.2	0	0	9.9
Culling, claws/legs **	0	2.4	5.8	0	0	7.8
Culling, fertility **	2.4	7.7	17.7	0	7.0	22.1
Cows with calving to first insemination interval, >70 d **	11.7	22.4	37.5	10.8	22.6	44.0
Cows with calving to latest insemination interval, >120 d	3.3	7.0	11.1	2.7	7.0	15.2
Incidence of paralysis and cramps **	0	3.6	8.4	0	1.6	10.8
Age at first calving (month)**	24.8	27.2	31.2	24.7	26.8	31.6

* Number of events per 100 animals at risk.

** Number of events per 100 animal-years at risk

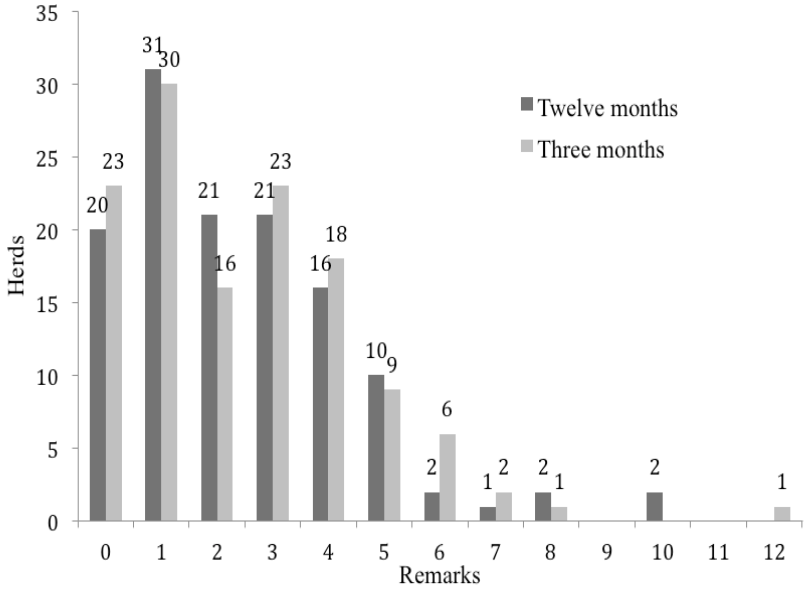


Figure 2. Number of herds per number of welfare remarks in 'Animal Welfare Signals', three (129 herds) and twelve (126 herds) months prior to an 'Ask the Cow' assessment. Herds among the 10% with the highest numbers in the welfare indicator were assigned with one welfare remark for that specific indicator.

4.2 Correlations

There were no strong statistically significant correlations between assessment points in 'Ask the Cow' and welfare indicators in 'Animal Welfare Signals' (data not shown). A significant (p-value <0.05) and relatively strong negative correlation (-0.5) was found between the welfare indicator of *mortality rate* and the assessment point of *rumen fill* in data from twelve months. Two significant and relatively strong negative correlations were found in data from three months; 1) *culling/mortality, first parity, 1-90 d after calving* and the assessment point of *rumen fill* (p-value <0.05 and correlation -0.5) and 2) the welfare indicator of *cows with calving to latest insemination interval >120 d* and the assessment point of *rumen fill* (p-value <0.05 and correlation -0.64). There were several more significant correlations, but they were low (range between 0.08-0.33 in twelve months and 0.17-0.32 in three months).

When the number of welfare remarks for every welfare indicator in all herds was summarized, about half the herds with a welfare remark within twelve months data also had a remark within three months data for that specific welfare indicator, and vice versa. However, in the welfare indicators of *heifers not bred, > 17 months of age, cows with calving to first insemination interval, >70 d, cows with calving to latest insemination interval, >120 d* and *age at first calving* there were a higher compliance between three and twelve months data.

4.3 Use of Welfare Indicators to Detect Herds With Poor Animal Welfare

A total of 29 (25.7%) and 17 (15.0%) of the remaining 113 herds were defined as having poor animal welfare according to 'Ask the Cow' (i.e. two or more welfare remarks) for project limits and currently applied limits, respectively. In Figure 3 the number of welfare remarks per herds is shown for project limits and currently applied limits.

Table 3 presents the sensitivity and specificity for each welfare indicator in 'Animal Welfare Signals' for twelve and three month tested against the gold standards based on project limits and current limits, respectively.

Sensitivity and specificity, positive predictive value and negative predictive value for the final combinations of welfare indicators from 'Animal Welfare Signals' in twelve or three months are presented in Table 4. Two of the test-kits had two combinations of welfare indicators with the same sensitivity and specificity; twelve months in project limits and three months in current limits. Table 5 displays sensitivity and specificity when the limits were changed into 80, 85 and 95% for each one of the welfare indicators separately.

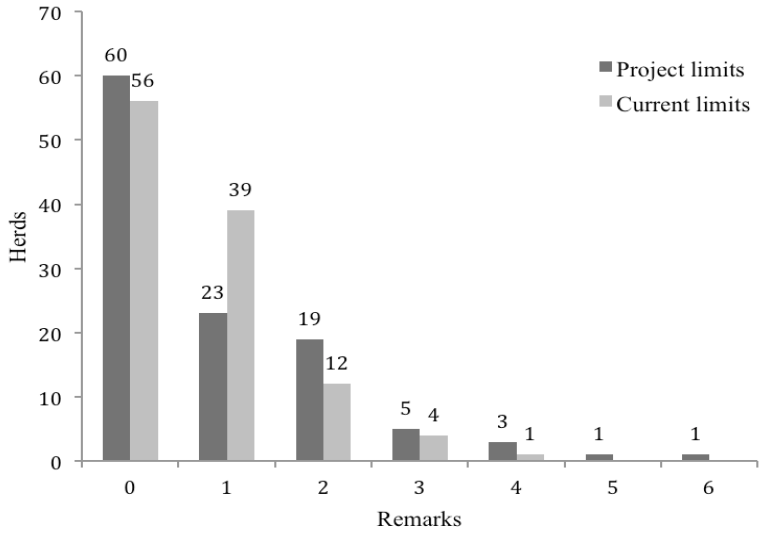


Figure 3. Number of welfare remarks per herd for the assessment points included in the gold standard for animal welfare; *lean, severely dirty* and *severe injuries* for all animal categories as well as *lameness* for cows. Herds with two or more welfare remarks were defined as having poor animal welfare. Numbers of remarks are shown for project limits and currently applied limits

Table 3. Sensitivity (Se) and specificity (Sp) for each welfare indicator in 'Animal Welfare Signals', when using them as tests to detect herds with poor animal welfare according to 'Ask the Cow'. For twelve (112 herds) and three (113 herds) months data in 'Animal Welfare Signals' as well as project limits and current limits in 'Ask the Cow'.

	Project limits				**Current limits**			
	Twelve months		**Three months**		**Twelve months**		**Three months**	
Welfare indicators	Se	Sp	Se	Sp	Se	Sp	Se	Sp
Calf mortality, 1-60 days *	10.3	90.4	13.3	90.4	17.6	91.6	16.7	90.5
Calf mortality, 2-6 months *	17.2	91.6	13.3	91.6	17.6	90.5	16.7	91.6
Young stock mortality, 6-15 months *	13.8	89.2	13.3	88.0	11.8	88.4	11.1	87.4
Culling/mortality, first parity, 1-90 d. after calving *	10.3	88.0	0	85.5	11.8	88.4	0	87.4
On-farm mortality **	6.9	90.4	10.0	92.8	5.9	90.5	11.1	92.6
Cows with diseases, total **	13.8	90.4	6.7	89.2	17.6	90.5	5.6	89.5
Low urea values (%)	0	89.2	3.3	89.2	5.9	91.6	0	89.5
Abnormal urea values (%)	6.9	88.0	10.0	91.6	5.9	88.4	5.6	90.5
Calving interval (months)	13.8	92.8	20.0	91.6	11.8	91.6	22.2	90.5
Incidence of calving difficulties **	17.2	94.0	10.0	91.6	17.6	92.6	16.7	92.6
Stillbirth, 0-24 hours *	24.1	94.0	10.0	94.0	23.5	91.6	11.1	93.7
Heifers not bred, > 17 months of age **	17.2	91.6	16.7	92.8	11.8	89.5	16.7	91.6
Culling total **	17.2	90.4	10.0	88.0	11.8	88.4	11.1	88.4
Incidence of other feed-related diseases **	3.4	86.7	13.3	89.2	0	87.4	11.1	88.4
Culling, udder diseases **	13.8	90.4	13.3	91.6	11.8	89.5	5.56	89.5
Incidence of mastitis treatments **	20.7	92.8	20.0	92.8	23.5	91.6	22.2	91.6
Somatic cell counts from test milking (1000 cells/ml)	10.3	90.4	16.7	95.2	11.8	90.5	16.7	93.7
Incidence of claw and leg disorders **	3.4	88.0	10.0	91.6	0	88.4	11.1	91.6
Culling, claws/legs **	0	86.7	0	88.0	0	88.4	0	89.5
Culling, fertility **	6.9	88.0	3.3	86.7	11.8	89.5	5.6	88.4
Cows with calving to first insemination interval, >70 d **	17.2	94.0	20.0	92.8	11.8	91.6	16.7	90.5
Calving - latest insemination, >120 d **	13.8	92.8	13.3	92.8	5.9	90.5	11.1	91.6
Incidence of paralysis and cramps **	0	85.5	3.3	86.7	0	87.4	5.6	88.4
Age at first calving (month)	20.7	92.8	26.7	90.4	23.5	91.6	33.3	89.5

* Number of events per 100 animals at risk.

** Number of events per 100 animal-years at risk

Table 4. Sensitivity (Se), specificity (Sp), positive predictive value (PPV) and negative predictive value (NPV) for the final test-kits of welfare indicators from 'Animal Welfare Signals' used as tests to identify herds with poor animal welfare according to the gold standard of 'Ask the Cow'. For twelve and three months as well as project limits and current limits, respectively.

	Se	**Sp**	**PPV**	**NPV**
Twelve months, project limits*	0.69	0.69	0.44	0.86
Three months, project limits**	0.73	0.68	0.45	0.88
Twelve months, current limits***	0.71	0.66	0.27	0.93
Three months, current limits****	0.72	0.71	0.32	0.93

*Welfare indicators included: *stillbirth, 0-24 hours; culling total; incidence of mastitis treatments; age at first calving* and *on-farm mortality* or *cows with calving to first insemination interval, >70 d*

**Welfare indicators included: *calving interval; culling udder diseases; incidence of claw and leg disorders; cows with calving to latest insemination interval, >120 d* and *age at first calving*

***Welfare indicators included: *low urea values (%); stillbirth, 0-24 hours; culling, udder diseases; incidence of mastitis treatments* and *age at first calving*

****Welfare indicators included: *calving interval; incidence of calving difficulties; stillbirth, 0-24 hours, age at first calving* and *calf mortality, 2-6 months* or *heifers not bred, > 17 months of age*

Table 5. Sensitivity (Se) and specificity (Sp) as well as positive predictive value (PPV) and negative predictive value (NPV) for the final test-kits of welfare indicators when one limit was changed. For twelve and three months as well as project limits and current limits respectively.

	Se	**Sp**	**PPV**	**NPV**
Twelve months, project limits*	0.76	0.63	0.42	0.88
Three months, project limits**	0.77	0.65	0.44	0.89
Twelve months, current limits***	0.82	0.61	0.28	0.95
Three months, current limits****	0.78	0.68	0.32	0.94

*Welfare indicators included: *stillbirth, 0-24 hours* (80%)*; culling total* (90%)*; incidence of mastitis treatments* (90%); *age at first calving* (90%) and *cows with calving to first insemination interval, >70 d* (90%)

**Welfare indicators included: *calving interval* (90%)*; culling udder diseases* (90%)*; incidence of claw and leg disorders* (90%)*; cows with calving to latest insemination interval, >120 d* (85%) and *age at first calving* (90%)

***Welfare indicators included: *low urea values* (%)(90%); *stillbirth, 0-24 hours* (90%); *culling, udder diseases* (90%); *incidence of mastitis treatments* (90%) and *age at first calving* (80%)

****Welfare indicators included: *calving interval* (90%); *incidence of calving difficulties* (90%); *stillbirth, 0-24 hours* (90%), *age at first calving* (90%) and *heifers not bred, > 17 months of age* (85%)

5. Discussion

5.1 Project and Currently Applied Limits

Results of this project show that there are differences between prevalences of the study and prevalences that are currently used as limits within 'Ask the Cow'. As the underlying data material was not the same this was somehow expected. It is probably to some extent due to the proportion of herds with loose-house systems, which differed between herds included in this project (70%) and in the study by Sandgren *et al* (2009) (33%). For natural reasons it is easier to detect lameness in cows in loose housed systems than in tied-up stalls, which can be an underlying reason for the higher project limit within the assessment point of *lameness* (Table 1). However it is important to take into account that for example an unclean floor may lead to increased lameness problems (Albright and Arave, 1997), which often is the case in loose housed systems. Compared to data material of currently applied limits, prevalences for the 90th percentile of the project were higher in the assessment points of *lameness*, *outwardly angled claws, mild injuries* and *severe injuries* in cows as well as of *injuries* in young stock. Otherwise prevalences of the project were lower compared to the data material of currently applied limits. These differences indicate that the limits for belonging to the 50th and 90th percentiles have changed and/or shifted for several of the assessment points. This could imply that currently applied limits are no longer representative for today's herds and stress the importance of regular monitoring of the results of assessments in a bench marking system like 'Ask the Cow' to secure that limits are representative and reflect reality. It is important to make sure that applied limits within 'Ask the Cow' can be trusted and based on results from this project it is probably necessary to validate them.

However, another important thing to consider is that herds from 'Ask the Cow' included by this project (from November 2009 until September 2012) were not randomly selected, which may affect how well they are representative for Swedish dairy herds. It is not known why herds included in the project performed an 'Ask the Cow'. There are several reasons for a herd to implement an 'Ask the Cow'-assessment (Thunberg and Winblad von Walter, 2011). Beyond those that are selected according to the rules of for example Swedish "KRAV" (KRAV, 2013), the reason for performing an 'Ask the Cow'-assessment is not known. The assessment can be requested by the herd manager, a veterinarian or an advisor with the purpose of for example improving animal welfare or learn more about the animal welfare status.

Limits practiced within 'Ask the Cow' are relative, i.e. a herd is compared to other participating herds and its welfare status depends on how good the other herds are. A positive aspect of this is that farmers can find out how their herds stand in relation to other herds, which could be a motivation to implement changes. Another method would be to use absolute limits, i.e. the herds' welfare status is rated based on what is acceptable for a certain level of welfare. An absolute limit could be less motivating from a farmer point of view, but it may be more reliable and sustainable in an animal welfare perspective. An absolute limit could

function as a more precise statement for improving animal welfare, which probably would be appreciated among consumers.

In 'Animal Welfare Signals', three divergences were detected when all welfare indicators were analysed for time periods of twelve and three months (Table 2). Two of them were within the ninetieth percentile and the welfare indicators *heifers not bred, >17 months of age* and *culling total*. Within these welfare indicators, limits were clearly lower in twelve months data compared to data within three months. On the other hand within the 10^{th} percentile and the welfare indicator of *culling total*, the limit was clearly higher in twelve months data compared to three months data. This implies that three-months and twelve-months data present different versions of the animal welfare situation at herd level within these specific welfare indicators. As within 'Ask the Cow' it is important to remember that figures from a shorter time interval display more how the situation appears at a specific moment (Winblad von Walter, personal communication, 2012). Therefore, it was of interest to investigate differences arising from different time intervals. Besides that, no distinctive differences in particular could be found in the distribution measurements between figures in twelve and three months.

5.2 Number of Welfare Remarks

The vast majority of assessed herds have one or less welfare remark within 'Ask the Cow' (Figure 1). Ten of totally 125 included herds have two or more welfare remarks for cows and are therefore considered as a case herd with poor animal welfare. Concerning young cattle and calves, only three and four herds had two or more welfare remarks of totally 117 and 121 herds included, respectively. However, a herd is classified as having poor animal welfare if they have two or more welfare remarks in total including all animal categories. This relatively small amount of case herds may be due to several things, were the first most credible reason could be that Swedish dairy herds keep a good animal welfare standard and therefore only a few herds are considered as case herds. However as earlier mentioned it would have been interesting to take into account why included herds performed an 'Ask the Cow'; for example if many of the herds voluntarily assigned to perform an 'Ask the Cow' the animal welfare situation could be different from herds assigned to an assessment because they do not fulfil the standards of for example KRAV. In addition, the fact that herds were not randomly selected could affect the outcome of the data material available in several different ways. For example the fact that one assessor have preformed twenty 'Ask the Cow'-assessments may affect the outcome of the data material. On the other hand, when assessors are properly trained the probability for validity is considered high in animal-based welfare measures (Knierim and Winckler, 2009 and Scott *et al*, 2001). All 'Ask the Cow' assessors are calibrated yearly (Thunberg and Winblad von Walter, 2011), but nevertheless, the human factor needs to be accounted for when any confounding factors that could affect the outcome from an 'Ask the Cow' assessment are discussed.

Based on register data from 'Animal Welfare Signals', a herd was given a welfare remark for each welfare indicator where they belong to the 90^{th} percentile. According to Figure 2, 78 of totally 126 included herds had two or more welfare remarks according to twelve months data

from ‘Animal Welfare Signals’. According to three months data, 53 of totally 129 included herds had two or more welfare remarks. This implies that a majority of herds included in the project have an animal welfare weakness according to twelve months data from ‘Animal Welfare Signals’. It could be assumed that herds with a good animal welfare standard are more accurate in reporting data presented in ‘Animal Welfare Signals’, but as many figures are reported automatically into the system that is not a credible reason for affecting results (Svensk Mjölk, 2011[3]). It can be considered as alarming that a majority of the herds included in the study had two or more welfare remarks within twelve months data, but it is important to have in mind that there could be several reasons for why a herd end up among the 10% worst herds. Especially for figures of three months data where for example a temporarily large amount of culled animals (excluded cows) could cause big changes in corresponding welfare indicators. The distribution of number of welfare remarks is more dispersed in ‘Animal Welfare Signals’ (range 0-12) than in ‘Ask the Cow’ (range 0-3) (Figure 1 and 2). This could be affected by that the number of assessment points and welfare indicators differs between ‘Ask the Cow’ and ‘Animal Welfare Signals’; theoretically the risk for a herd being classified as having poor animal welfare could be higher in ‘Animal Welfare Signals’ were there are more points verified.

5.3 Diagnostic Tests

Diagnostic tests were used with the aim of creating a test-kit of welfare indicators that could be used for identifying herds with good or poor animal welfare (Sandgren *et al*, 2009). According to figures presented in Table 3, sensitivity and specificity are similar for each welfare indicator in ‘Animal Welfare Signals’ for twelve months and three months within project limits. The same applies within current applied limits of ‘Ask the Cow’ as well as when results of sensitivity and specificity are compared across limits. However, for some welfare indicators there were larger differences in sensitivity across project and current limits in three and twelve months. For example, three of the welfare indicators had a large difference in sensitivity between project and current limits within both three and twelve months; *culling/mortality, first parity 1-90 d after calving, stillbirth 0-24 hours* and *incidence of other feed-related diseases*. However, it is difficult to draw any conclusions out of these results, as the difference is not tested for significance. Sensitivity is the ratio of herds with poor animal welfare that test positive (Thrusfield, 2005). Overall, sensitivity is low for all included welfare indicators (range 0-33.3) and this motivates the fact that one specific indicator is not enough for detecting herds at risk of having poor animal welfare.

As can be seen in Table 4, sensitivity and specificity had similar results in all four test-kits. The highest found sensitivity was 0.73 (Table 4). The final value was thereby higher than if welfare remarks would be picked out by chance (>50%) and about as high as in the study by Sandgren *et al* (2009) were the final test-kit resulted in a sensitivity of 0.77. Sandgren *et al*, 2009 found a specificity of 0.91, while the highest found specificity in the final test-kit of this project was lower; 0.71. Sandgren *et al* (2009) included a total of 55 dairy herds and in this project the minimum number of dairy herds analysed for epidemiological measurements were 112. A consequence of more data available would have been a more reliable result, which Sandgren *et al* also emphasizes. Despite that, sensitivity and specificity were lower within this

project compared to results from the study by Sandgren et al (2009). However, it is important to remember that society changes and that a lot has happened within the dairy sector the last couple of years. For example, due to a tough economic market many herds have been forced to close down (DN, 2013) and this has resulted in less but larger herds (Svensk Mjölk, 2012[3]). 'Ask the Cow'-assessments cost money and therefore bad economy may result in less 'Ask the Cow'-assessments performed and/or that the majority of herds that perform them are those who can afford it. In the study by Sandgren *et al* (2009), the proportion of case herds with poor animal welfare is similar (23.6%) to currently applied limits (25.7%), while the proportion is smaller within project limits (15.0%). This difference may also affect the results as a smaller proportion of case herds could lead to less reliable numbers.

None of the four final test-kits presented in Table 4 are consistent with each other. However, the welfare indicator of *age at first calving* is represented in all models and can therefore be assumed as important when detecting herds at risk of having poor animal welfare. In the study by Sandgren *et al* (2009), *age at first calving* was not represented in any of the found test-kits. This difference may indicate that something has changed that lead to *age at first calving* now being specifically important. Within twelve months, the final test-kit of project limits have three welfare indicators in common with currently applied limits; *incidence of mastitis, stillbirth, 0-24 hours* and *age at first calving.* The final test-kits within three months had two welfare indicators in common; *calving interval* and *age at first calving.* In two of the test-kits, two combinations of welfare indicators with the same sensitivity and specificity were found. The final test-kit in the study by Sandgren *et al* (2009) was included by the welfare indicators of *calf, young stock* and *cow mortality* as well as *heifers not bred >17 months* and *cows with late on-going artificial insemination (Als), >120 d* (or "*calving - latest insemination, >120 d"*). All but *cow* and *young stock mortality* are representative in any of the final test-kits of this project and the rest are represented in any of the test-kit of this project, but then spread between three of the four different test-models. Of this, one could conclude that it is difficult to found a specific set of welfare indicators that are assumed as important when detecting herds at risk of having poor animal welfare. However, to not settle with only a few welfare indicators could be one solution. The test-kit was similar to parallel interpretation of tests, i.e. a herd would be counted as having poor animal welfare if it had a remark in any of the welfare indicators included (Thrusfield, 2005). To avoid false positive and false negative values, it is desirable to have a high sensitivity without lowering specificity too much. There could be several reasons for the large difference in welfare indicators in results between this project and the study by Sandgren *et al* (2009). First of all and as earlier mentioned, underlying data material in the study by Sandgren *et al* (2009) was about half of the data material obtainable in this project and this could affect results given by both studies. Another important point to take into consideration is that included assessment points of the gold standard differ between this project and the study by Sandgren *et al* (2009). Although the same limit was practiced in both studies (a herd was given a welfare remark if among the 10% worst and two or more welfare remarks equals to poor animal welfare), included assessment point from 'Ask the Cow' of this project were only those available (for example not *rising behaviour*) and assumed as extremes regarding poor animal welfare.

Within parallel testing sensitivity will increase and specificity will decrease. This will also lead to an increased negative predictive value. Compared with results in the study by Sandgren *et al* (2009), all positive predictive values were considered relatively low in all four of the test-kits (a range between 0.27-0.45 compared to 0.5) (Table 4). The positive predictive values of currently applied limits in 'Ask the Cow' were distinctively lower than in project limits; 0.27 and 0.32 compared to 0.44 and 0.45 respectively. If the positive predictive value is low, there is a higher risk for herds with a good animal welfare standard being incorrectly classified as having poor animal welfare instead. This is probably not preferred among farmers as the risk of being wrongly accused for having poor animal welfare increases. However, it is important to remember that positive predictive values are affected by sensitivity, specificity and prevalence; the last two in particular (Thrusfield, 2005). If prevalence decreases, so does the positive predictive value. Negative predictive values found were considered high (a range between 0.86-0.93), which means that the proportion of herds that test negatively and truly have a good animal welfare was high (Appendix 4) (Thrusfield, 2005). This could be of interest among consumers and used as an argument for trusting the system.

Finally it was interesting to look further into if sensitivity could be improved by changing the limits of one welfare indicator included in the test-kit at a time. In two out of four test-kits, sensitivity was improved when the limit for one welfare indicator was changed into 80% instead of 90% (Table 5). These were *stillbirth, 0-24 months* for project limits and *age at first calving* for current limits in twelve months. Within the test-kits for project and current limits in three months, sensitivity was improved when two limits were changed into 85% for the welfare indicators of *cows with calving to latest insemination interval, >120 d* and *heifers not bred, > 17 months* instead. However, the difference in sensitivity when the limit was changed compared to when all were analysed according to 90% was not particularly distinctive.

5.4 Correlations

Statistically significant and relatively strong negative correlations were found between three different combinations of welfare indicators and assessment points. The welfare indicator of *rumen fill* is found within all three of the significant and relatively strong correlations. Only 16 herds have assessed rumen fill (Table 1), but since there is a correlation between *mortality rate* and two fertility traits it probably should be taken more seriously when cows are assessed within 'Ask the Cow'. It would be interesting to know why the assessors have chosen to neglect the assessment point of *rumen fill* since it is easy to assess and tells a lot about the welfare status of the animal (Hulsen, 2011). Probably one reason is that *rumen fill* is voluntary for the assessor to measure (Svensk Mjölk, n.y.[1]). Other significant correlations found were considered low (range between 0.08-0.33 in twelve months and 0.17-0.32 in three months) and not interesting to proceed with.

Results show that about half of the welfare remarks are equally present in both twelve and three months in almost all welfare indicators. This indicates that welfare remarks from 'Animal Welfare Signals' in twelve and three months are relatively equivalent in most of the welfare indicators. However, in five of the welfare indicators (*on farm mortality; heifers not*

bred, > 17 months of age; cows with calving to first insemination interval, >70; cows with calving to latest insemination interval, >120 and *age at first calving*) there is a larger difference in distribution between twelve and three month data. This could partly be due to the characteristics of the indicator; if it is affected by time of year for collection of data or if data was collected after a specifically tough time-period. However, within this material it was not possible to view a notable difference that depends on the time period.

6. Conclusions

Result of this study shows that according to how the reality changes on herd level, currently applied limits of ‘Ask the Cow’ need to be followed up and adjusted for. All found test-kits identify herds with a good animal welfare to a high degree. However, there is a higher risk of herds with good animal welfare being incorrectly classified as having poor animal welfare. Since the final test-kits were not particularly consistent with each other, it is recommended to apply several welfare indicators when identifying herds at risk of having poor animal welfare. However, the welfare indicator of *age at first calving* was found in all final test-kits and can therefore be considered as particularly interesting. ‘Ask the Cow’ and ‘Animal Welfare Signals’ do reflect the same view of animal welfare, but as result reviles it is recommended to use both of them when assessing animal welfare at herd level. Animal welfare is a challenging concept and therefore it could be discussed if ‘Ask the Cow’ and ‘Animal Welfare Signals’ have to reflect the same view of animal welfare at all; if they simply complement each other instead.

7. References

Albright, J.L. and Arave, C.W. (1997) *The Behaviour of Cattle.* Cab International, Wallingford, UK. ISBN 0-85199-196-3. Page 31.

Bareille, N., Beaudeau, F., Billon, S., Robert, A. and Faverdin, P. (2003) *Effect on health disorders on feed intake and milk production in dairy cows.* Livestock Production Science 83, page 53–62, France

Blokhuis, H.J., Jones, R.B., Geers, R., Miele, M. and Veisser, I. (2003) *Measuring and monitoring animal welfare: transparency in the food product quality chain.* UFAW and Animal Welfare 2003 12: 445-455

Broom, D. M. (1991) *Animal Welfare: concepts and measurements.* Journal of Animal Science, 69: 4167-4175

Broom, D. M. (1988) Les concepts de stress et de bien-être. Rec. Med. Vet. 164:715

Council Directive concerning the protection of animals kept for farming purposes (1998-06-01). (98/58/EC)

Council Directive of laying down minimum standards for the protection of calves (2008-12-18). (2008/119/EC)

Dagens Nyheter (DN). *Antalet svenska bönder minskar dramatiskt.* [online] (2012-05-27) Available: http://www.dn.se/ekonomi/antalet-svenska-bonder-minskar-dramatiskt [2013-05-29]

European convention for the protection of animals kept for farming purposes (1978-11-17). (L 323)

Farm Animal Welfare Council (FAWC). *Five Freedoms.* [online] (2009-04-16) Available: http://www.fawc.org.uk/freedoms.htm [2013-04-16]

Fraser, A.F. (1989) *Animal Welfare practice: Primary Factors and Objectives.* Applied Animal Behaviour Science, The Netherlands. 22: 159-176

Fraser, D., Weary, D. M., Pajor, E. A. and Milligan, B. N. (1997) *A scientific conception of animal welfare that reflects ethical concerns.* Centre for Food and Animal Research, Canada. Animal Welfare. 6: 187-205. *Agriculture and Agri-Food Canada, Ottawa KIA OC6. Canada*

Fraser, D. (2008) *Understanding Animal Welfare: The Science in its Cultural Context.* UFAW, ISBN: 978-1-4051-3695-2, page: 233

Forkman, B. and Keeling, L. (2009) *Assessment of Animal Welfare Measures for Dairy Cattle, Beef Bulls and Veal Calves.* Welfare Quality® Reports No 11, Uppsala. ISBN 1-902647-81-5. Page: zero (contents).

Hulsen, J. (2011) *Kosignaler - En praktisk bok om mjölkföretagande med kon i fokus.* Roodbont Förlag (Zutphen) and Vetvice (Bergen). ISBN 978-90-810974-3-7. Page: II, 6-9, 21, 52 and 77.

Knierim, U. and Winckler, C. (2009) *On-farm welfare assessment in cattle: validity, reliability and feasibility issues and future perspectives in special regard to the Welfare Quality® approach.* Animal Welfare 2009, 18: 451-458. ISSN 0962-7286.

KRAV. *5.10.4. Uppföljning för mjölkkobesättningar*. [online] (2013) Available: http://www.krav.se/node/36496#booknode-36503 [2013-07-19]

Mellor, D. J., Patterson-Kane, E. and Stafford, K. J. (2009) *The Science of Animal Welfare*. The Universities Federation for Animal Welfare, Wiley-Blackwell, ISBN 978-1-4051-3495-8. Page 4.

Mörk, M. *Välfärdsindikatorer i Kodatabasen.* Forskning Special, nr 4. [online] (2011-03-16) Available: http://svenskmjolk.se/Global/Dokument/Dokumentarkiv/Forskning/Forskning%20Special/Forskning%20Special%202011_04%20V%C3%A4lf%C3%A4rdsindikatorer%20i%20kodatabasen.pdf [2013-06-11]

Rushen, J., de Passillé, AM., von Keyserlingk, MAG. and Weary, DM. (2008) *The Welfare of Cattle*. Volume 5. The Netherlands, Springer. ISBN 978-1-4020-6557-6. Page preface, 7, 15, 43, 58 and 77.

Sandgren, CH., Lindberg, A. and Keeling, LJ. (2009) *Using a national dairy database to identify herds with poor welfare.* Animal Welfare 18:523-532.

Scott, E.M., Nolan, A.M. and Fitzpatrick, J.L. (2001) *Conceptual and methodological issues related to welfare assessment: A framework for measurement.* Acta Agriculturae Scandinavica Section A, Animal Science Supplementum *30*: 5-10

Stata® Corporation, College station, TX, USA

Statens jordbruksverks föreskrifter och allmänna råd om djurhållning inom lantbruket m.m. (2010). Jönköping. (SJVFS 2010:15, L100)

Svensk Mjölk:

1. *Stabilt antal kor i Kokontrollen*. [online] (2011-01-21) Available: http://svenskmjolk.se/Mjolkgarden/Mjolkgarden/2011/Stabilt-antal-kor-i-kokontrollen/ [2012-06-11]
2. *Att använda Signaler Djurvälfärd.* (2011-01-27)
3. *Signaler Djurvälfärd – Formelsamling.* (2011-01-27)

Svensk Mjölk. No year (n.y.):

1. Bedömningsprotokoll. Fråga Kon.
2. Signaler Djurvälfärd och Fråga Kon.

Svensk Mjölk:

1. *Kokontrollen basen för beslut.* [online] (2012-01-11) Available: http://svenskmjolk.se/Mjolkgarden/Service--radgivning/Kokontrollen/ [2012-06-11]
2. *Större, men färre besättningar i Kokontrollen.* [online] (2012-09-18) Available: http://www.svenskmjolk.se/Mjolkgarden/Tidningen-Husdjur/Nyheter/2012/Storre-men-farre-besattningar-i-Kokontrollen/#.UaXNAaWsFo4 [2013-05-29]

The animal protection ordinance (1988). Stockholm. (SFS 1988:539).

The animal welfare act (1988). Stockholm. (SFS 1988:534).

Thrusfield, M. (2005) *Veterinary Epidemiology – third edition.* Blackwell Science Ltd, Great Britain. ISBN 978-06320-6397-0. Page: 159, 312-317 and 323.

Thunberg, K. and Winblad von Walter, L. (2011-05-06). *Fråga kon – praktiskt i djurbesättning.* Svensk Mjölk.

Welfare Quality®:

1. *Welfare Quality Video.* [online] (2004) Available: http://www.animalfarmlife.eu/dvd/ [2013-04-18]
2. *Welfare Quality®: Science and society improving animal welfare in the food quality chain.* [online] (2004) Available: http://www.welfarequality.net/everyone/26536/5/0/22 [2013-04-18]

Welfare Quality® (2009). *Welfare Quality® assessment protocol for cattle.* Welfare Quality® Consortium, Lelystad, Netherlands.

Winckler, C., Algers, B., Boivin, X., Butterworth, A., Canali, E., De Rosa, G., Hesse, N., Keeling, L., Knierim, U., Laister, S., Leach, K., Milard, F., Minero, M., Napolitano, F., Schmied, C., Schul- ze Westerath, H., Waiblinger, S., Wemelsfelder, F., Whay, H.R., Windschnurer, I. and Zucca, D. (2007) *Assuring animal welfare: from societal concerns to implementation. Dairy Cows.* Programme Second Welfare Quality® stakeholder conference Berlin. ISBN number 978-90-78240-02-0

7.1 Unpublished Material

Mörk, Marie. Personal communication. December 2012 and May 2013.

Roth, Anki. Personal communication: Lecture in Cow Signals. 2013-04-11.

Winblad von Walter, Louise. Personal communication. 2012-09-24.

Winblad von Walter, Louise. Personal communication: email. 2013-05-20.

8. Appendix

Appendix 1

Welfare principles with associating criterias and descriptions from the Welfare Quality® project (Welfare Quality®, 2009):

"Welfare principles	*Welfare criteria*	
Good feeding	*1*	*Absence of prolonged hunger*
	2	*Absence of prolonged thirst*
Good housing	*3*	*Comfort around resting*
	4	*Thermal comfort*
	5	*Ease of movement*
Good health	*6*	*Absence of injuries*
	7	*Absence of disease*
	8	*Absence of pain induced by management procedures*
Appropriate behaviour	*9*	*Expression of social behaviours*
	10	*Expression of other behaviours*
	11	*Good human-animal relationship*
	12	*Positive emotional state*

The principles and criteria that are the basis for the Welfare Quality® assessment protocols. More detailed definitions of welfare criteria are described below.

1. *Animals should not suffer from prolonged hunger, i.e. they should have a suitable and appropriate diet.*
2. *Animals should not suffer from prolonged thirst, i.e. they should have a sufficient and accessible water supply.*
3. *Animals should have comfort when they are resting.*
4. *Animals should have thermal comfort, i.e. they should neither be too hot nor too cold.*
5. *Animals should have enough space to be able to move around freely.*
6. *Animals should be free of injuries, e.g. skin damage and locomotory disorders.*
7. *Animals should be free from disease, i.e. animal unit managers should maintain high standards of hygiene and care.*
8. *Animals should not suffer pain induced by inappropriate management, handling, slaughter, or surgical procedures (e.g. castration, dehorning).*
9. *Animals should be able to express normal, non-harmful, social behaviours (e.g. grooming).*
10. *Animals should be able to express other normal behaviours, i.e. it should be possible to express species-specific natural behaviours such as foraging.*
11. *Animals should be handled well in all situations, i.e. handlers should promote good human-animal relationships.*
12. *Negative emotions such as fear, distress, frustration or apathy should be avoided whereas positive emotions such as security or contentment should be promoted."*

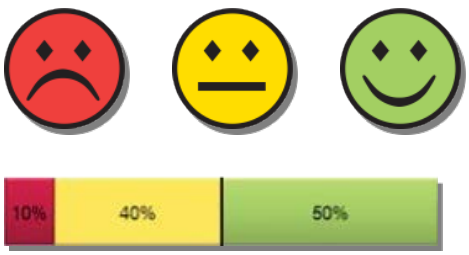
Appendix 2

Welfare indicators from 'Animal Welfare Signals' and smileys that represent animal welfare on herd level as well as the 10th, 50th and 90th percentile for each colour respectively.

Areas	Welfare indicators
Calves	Calf mortality, 1-60 days*
Calves	Calf mortality, 2-6 months*
Heifers	Young stock mortality, 6-15 months*
Heifers	Culling/mortality, first parity, 1-90 d after calving*
Sustainability	On-farm mortality**
Diseases	Cows with diseases, total**
Feed balance	Low urea values (%)
Feed balance	Abnormal urea values (%)
Monitoring	Calving interval (months)
Calvings	Incidence of calving difficulties**
Calvings	Stillbirth, 0-24 hours*
Heifers	Heifers not bred, > 17 months**
Sustainability	Culling total*
Feed balance	Other feed-related diseases*
Sustainability	Culling, udder diseases*
Diseases	Mastitis treatments*
Diseases	Somatic cell counts from test milking (1000 cells/ml)
Sustainability	Incidence of claw and leg disorders**
Sustainability	Culling, claws/legs**
Monitoring	Culling, fertility**
Monitoring	Cows with calving to first insemination interval, >70 d**
Monitoring	Calving - latest insemination, >120 d**
Sustainability	Incidence of paralysis and cramps*
Heifers	Age at first calving (months)

* Number of events per 100 animals at risk.

** Number of events per 100 animal-years at risk



Appendix 3

Assessment points within each animal category from 'Ask the Cow' and the flower used for presenting results given by the assessment of cows.

Table 1. Assessment points for cows

Assessment points	Rating scale	Scale of poor animal welfare
Rising behaviour	1- 4 (or 2 – 5*)	3 and 4 (or 4 and 5*)
Body condition	2 - 4	2 (lean) and 4 (fat)
Cleanliness	2 - 4	3 and 4
Dusty neck**	No remarks or dusty neck	Dusty neck
Scuffed neck**	No remarks or scuffed neck	Scuffed neck
Claw length***	No remarks, claw length > 95 mm	Claw length
Outwardly angled claws***	No remarks or deflected back claws	Outwardly angled claws
Lameness	0, 2 or 4	2 and 4
Injuries	2 - 4	3 and 4
Rumen fill	1, 2 or 4 (or 1-3*)	2 and 4 (or 2-3*)
Injury placement****	Carpus, front, hock, back part or others	

* Scale practiced within 'Ask the Cow'

** Can show both at the same time, counted for as one assessment point; i.e. coat

*** Can show both at the same time, counted for as one assessment point; i.e. claws

**** Not included

Table 2. Assessment points for young stock

Assessment points	Rating scale	Scale of poor animal welfare
Body condition	2 - 4	2 (lean) and 4 (fat)
Cleanliness	2 - 4	3 and 4
Injuries	2 - 4	3 and 4
Injury placement*	Carpus, front, hock, back part or others	-
Health	No remarks, diarrhoea, forced breathing, lameness or indolence	Diarrhoea, forced breathing, lameness or indolence

* Not included in project

Table 3. Assessment points for calves

Assessment points	Rating scale	Scale of poor animal welfare
Body condition	2 - 4	2 (lean) and 4 (fat)
Cleanliness	2 - 4	3 and 4
Injuries	2 - 4	3 and 4
Injury placement*	Carpus, front, hock, back part or others	-
Health	No remarks, diarrhoea, forced breathing, lameness or indolence	Diarrhoea, forced breathing, lameness or indolence

* Not included in project

Appendix 4

Diagnostic tests; test result and disease ratio.

	With disease (sick)	**Without disease (healthy)**	**Total**
Positive test	a. True positive	b. False positive	a + b
Negative test	c. False negative	d. True negative	c + d
Total	a + c	b + d	

Sensitivity (Se) = a / (a + c)

Specificity (Sp) = d / (b + d)

Positive predictive value (PPV) = a / (a + b)

Negative predictive value (NPV) = d / (c + d)