© 2009 Universities Federation for Animal Welfare
The Old School, Brewhouse Hill, Wheathampstead,
Hertfordshire AL4 8AN, UK

Animal Welfare 2009, 18: 523-532
ISSN 0962-7286

Using a national dairy database to identify herds with poor welfare

CH Sandgren*†, A Lindberg†‡ and LJ Keeling§

† Swedish Dairy Association, PO Box 210, SE 101 24, Stockholm, Sweden
‡ National Veterinary Institute, Department of Disease Control, SE 751 89, Uppsala, Sweden
§ Department of Animal Environment and Health, Swedish University of Agricultural Sciences, Box 7068, SE 750 07, Uppsala, Sweden
* Contact for correspondence and requests for reprints: charlotte.sandgren@svenskmjolk.se

Abstract

This research project was carried out by the Swedish Dairy Association as part of the development of a 'Scheme for Animal Welfare'. As a first step, on-farm, animal-based measurements on calves, young stock and cows in 55 randomly selected herds were performed. Nine animal-based measurements formed the basis for a classification of welfare at the herd level: cleanliness and body condition in calves, cows and young stock, in combination with lameness, injuries/inflammations and rising behaviour which were recorded for cows only. The threshold (gold standard) for being a case herd with poor welfare, was a score lying among the worst 10% on two or more of the nine welfare measurements. Thirteen of the 55 herds were cases fitting this criterion. As a second step, 65 potential welfare indicators from seven different focus areas in a pre-collected data register were identified by expert opinion. In the final step, the extent to which suggested potential welfare indicators predicted farms' risk of having poor welfare according to the definition was assessed. The final set of welfare indicators, taken from the national dairy database, included two fertility measures and calf mortality, and it correctly classified 77% of the herds, with a sensitivity of 0.62. The inclusion of cow and young stock mortality led to it correctly classifying 76% of the herds with a sensitivity of 0.77. We propose that this approach could be useful in helping to allocate advisory services to farms at risk of poor welfare.

Keywords: *animal based, animal welfare, epidemiology, national dairy database, pre-collected data, risk*

Introduction

Traditionally, resource-based measures such as floor type, space allocation, etc have been included in welfare legislation and schemes (Swedish Animal Welfare Act 1988; Bartussek 1999). However, during the past ten years, research into methods for evaluating animal welfare at the herd level has increased tremendously (Webster & Main 2003; Winckler *et al* 2007). In 2003, a list of parameters regarded as being sufficiently well documented to be included in a cattle welfare measuring system was published (Winckler *et al* 2003) and today, there is considerable agreement to use mainly animal-based measures, such as lameness, body condition and cleanliness in assessments of animal welfare in a dairy herd (Keeling & Veissier 2005; Veissier & Evans 2007). The herd-to-herd variation in animal-based measurements is considerable as these parameters are affected both by management and production system (Rousing 2003; Whay *et al* 2003). But, there are concerns about whether welfare assessment systems containing animal-based measures can feasibly be implemented in practice if a large number of herds are to be monitored on a regular basis. The Nordic cow databases contain a comprehensive list of records (fertility, disease, mortality, culling reasons, milk quality and milk production, etc) and have the potential to integrate significant levels of scientific knowledge in the field of animal welfare with that in previously existing monitoring systems (Philipsson *et al* 2003; Fraser 2004).

Given this as a backdrop, the Swedish Dairy Association was commissioned to develop a 'Scheme for Animal Welfare', which could be used by all Swedish dairy farmers. The aim being to increase the well-being of the animals, to strengthen the profitability of Swedish dairy herds and safeguard consumer trust. Part of the project was carried out in co-operation with the Danish Dairy Federation.

Three areas of research were identified: can dairy herd welfare be estimated using information in existing databases; can the farmer be stimulated to address welfare issues in his/her herd and do farmers that have herds with good welfare have a better farm economy. This study seeks to address the first of these questions by investigating whether dairy herds with poor welfare, determined by on-farm, animal-based welfare assessments, can be reliably predicted using a set of welfare indicators from a pre-collected data register.

Table 1 Animal-based welfare measurements and cut-off levels used in a field study on dairy cow welfare including 55 randomly selected Swedish dairy farms.

Animal-based welfare measurement	Definition of cut-offs (resulting in a remark)		
	Calves 0–6 months	Young stock 6 months–calving	Dairy cows
Cleanliness	Areas of old manure	Areas of old manure	Areas of old manure
Body condition	Concave lumbar back region	Concave lumbar back region	Body condition score ≤ 2 (Gillund *et al* 1999)
Injuries and inflammations			Summarised extensions of injuries or inflammations > 4 cm
Lameness			Lame to severely lame (Sprecher *et al* 1997)
Rising behaviour			Normal sequence with difficulties to finalise the rising or abnormal rising behaviour, eg dog-sitting, crawling backwards before rising etc (Chaplin & Munksgaard 2001)
Avoidance distance			Moves backwards when approached or avoids assessor completely (Rousing & Waiblinger 2004)

Materials and methods

Study herds

Data from 55 dairy herds, affiliated to the Official Milk Recording Scheme, were included in this study. Eighteen of these were loose housed in cubicle stalls and 37 were housed in tie stalls. Six of the herds had Swedish Red and White (SRB) cattle. Fourteen had Swedish Holstein Friesian cattle (SHF) and 35 had mixed breeds. Twenty-nine of the herds had 20–49 cows, 20 had 50–99 cows and six had 100–415.

Animal-based measurements as a basis for the welfare gold standard

Ten animal-based welfare measurements were collected on-farm. These were cleanliness and body condition in calves, cows and young stock, in combination with lameness, injuries/inflammations, rising behaviour and avoidance distance which were recorded only for cows (Table 1). Each farm was assessed twice in 2005, the first occasion from March to the middle of June and the second from October to December. At each visit two assessors, independently and without communicating, performed an assessment. At the second visit one of the previous assessors was replaced, so that each farm was visited by three assessors, one of whom had visited the farm twice.

Herd-level estimates of the animal-based measurements were obtained by applying the cut-off levels seen in Table 1 and by calculating the proportion of animals within each age group that exceeded the cut-off. The welfare 'gold standard' ie the definition of welfare status against which the performance of potential welfare indicators in the database was to be evaluated, was based on the number of animal-based measurements where a herd scored among the worst 10%. This led to the herd being attributed a 'remark' for that measure and any herd with two or more remarks was considered poor in terms of welfare.

Data sources for potential welfare indicators

The database of the Swedish national dairy recording system (hereafter called the Swedish Cattle Database [SCD]) was used as the main source of potential welfare indicators for this study. The SCD includes information on, eg fertility, genetics, diseases, mortality, including culling reasons, production and slaughterhouse registrations as well as demographic data (Emanuelson 1988; Olsson *et al* 2001).

The database is restricted to those herds affiliated to the Official Milk Recording Scheme (approximately 80% of Swedish dairy herds). Additional data on cattle mortality (enabling identification of euthanised and fallen stock) was retrieved from the Board of Agriculture, where records of all Swedish cattle are kept, in accordance with EU Directive 1760/2000. This was done in order to investigate the usefulness of such data in the event of a herd not being affiliated to the Official Milk Recording System and also because at the time of the study the SCD did not separate 'euthanasia' from 'fallen stock'.

Choice of potential welfare indicators from the data sources

Seven different focus areas reflecting important components of animal welfare and covering the complete lifespan of a dairy cow were suggested by a group of experts on quality programmes, marketing, dairy farm economy and animal health, within the Swedish Dairy Association and Danish Cattle Federation. The seven areas were: management; calves and young stock; survival/intensity of production; feeding; udder health; claw and leg health and drug use.

In a second step, a total of 65 potential welfare indicators from the SCD and the Board of Agriculture data register that could be expected, in varying degrees, to reflect the welfare aspects of the focus areas, were identified by a group consisting of twelve national experts in animal health,

Table 2 Potential welfare indicators from pre-registered data, covering seven focus areas and evaluated with respect to their association with welfare as defined by animal-based measurements, on-farm, in a study involving 55 randomly selected Swedish dairy farms.

Focus area	Potential welfare indicator[1]
Management	Age at first calving (months), average calving interval; CV[2] calving interval; % heat detection; % culling for fertility reasons; % cows with late beginning of AIs, > 70 days; % cows with late ongoing AIs, > 120 days; incidence of fertility treatments; number of inseminations per series
Calves and young stock	% calving difficulties; % stillbirths; calf mortality 1–60 days; calf mortality 1–90 days; calf mortality 2–6 months; % live born calves per AI series; average 305 days milk yield (kg), primiparous cows; CV 305 days milk yield primiparous cows; average kg milk per day 2nd–3rd lactation month, primiparous cows; CV average milk per day (kg) 2nd-3rd lactation month, primiparous cows; young stock mortality, 6–15 months; % survival, primiparous cows early lactation (3 months); average number of inseminations per heifer; incidence of disease treatments in young stock 6–15 months; % heifers not bred > 17 months; growth per day, heifers, chest girth at first calving
Survival/intensity of production	% cow mortality; % cows ≥2nd lactation; % cows ≥ 3rd lactation; % survival, early in lactation, cows (3 months); average kg 305 days milk yield, cows; CV 305 days milk yield, cows; average kg milk per day 2nd–3rd months, cows; CV average milk per day (kg) 2nd–3rd lactation month, cows; % no remarks at slaughter; % culling related to diseases[3]; % voluntary culling[3], % culling due to diseases other than mastitis, fertility or claw and leg disorders; % culling, overall, replacement rate
Feeding	% cows in good condition at slaughter; % cows < 3% fat content at three consecutive milk recordings; % cows < 3% protein content at three consecutive milk recordings; incidence of paralytic or tetanic conditions; incidence of other feed-related diseases including gastrointestinal disturbances; prevalence of cows with high levels of urea[4]; prevalence of cows with with high levels of urea[4]; prevalence of cows with low levels of urea[4]; prevalence of cows with urea remarks (high and low)[4]
Udder health	Incidence of mastitis treatment; % cows with no mastitis treatment; % culling for udder health reasons; survival of cows with mastitis treatment > 3 months; prevalence of cows with chronically high cell counts[5]; prevalence of cows with low cell counts[5]; % incidence risk, udder infections[6]; incidence rate, udder infections[6]; % cows with new chronic infections[5]; bulk milk somatic cell count (BMSCC), milk recording; BMSCC, delivered milk
Claw and leg health	% incidence claw and leg diseases; % culling for claw and leg disorders
Drug use	% cows with no veterinary treatment

[1] Principles for calculations (in Swedish) may be retrieved from the author upon request.
[2] Coefficient of variation.
[3] ± cases with code 'other reasons'.
[4] According to definition by Gustafsson (1993).[5] Brolund (1985).
[6] Based on a SCC > 150 (at risk = all heifers, and cows with ≥ 3 milk recordings with SCC < 150).

welfare, production and epidemiology (Table 2). The values of the potential welfare indicators were calculated for the period Jan–Dec 2005 for the 55 farms, ie over the same period of time as the on-farm data were collected.

Data analysis

Variable reduction

The aim of the analyses was not to study causality, nor to determine exact relationships between animal-based measurements and potential welfare indicators. Rather, the aim was to identify a set of pre-recorded welfare indicators that could, in combination, be used to identify herds at risk of poor welfare. Given this aim, it was necessary to reduce the initial set of potential welfare indicators suggested by the experts. Therefore, as a first step, univariable associations between all 65 potential welfare indicators and each of nine animal-based measurements were screened using linear regression. Welfare indicators with an association significant at $P < 0.05$ were then taken forward to a multivariable reduction step, using the same methodology. By including the second step we added a stronger requirement that the potential indicator should show a significant multivariable association with one or more of the animal-based measurements, in order to be taken further. Consequently, only indicators that were significantly ($P < 0.05$) associated with one or more animal-based measures in this multivariable context were considered to be candidates for the final set of welfare indicators. All statistical analyses were performed using the software Stata® version 10 (Stata Corp, College Station, TX, USA).

Selection of final set of welfare indicators (the 'test tool')

We chose to regard each potential welfare indicator as a 'diagnostic test', ie a tool that distinguishes between two different statuses; eg sick vs healthy; pregnant vs not pregnant; good vs poor welfare etc. As with any diagnostic test measured on a continuous scale, cut-off levels needed to be identified as these are the points at which the 'test' would be regarded as being 'positive'. Since the choice of cut-off level affects 'test' performance, measured as sensitivity (the probability of correctly identifying case herds) and specificity (the probability of correctly identifying herds that are not cases), we evaluated each potential welfare indicator at three different levels. The cut-off levels evaluated were the 80th, 90th and 95th percentile for welfare indicators that were positively associated with the animal-based measurements and the 20th, 10th and 5th percentile where the association was negative. In this way, all potential welfare indicators were dichotomised (positive or not) into three different 'tests' and the sensitivity and specificity for each 'test' in identifying a herd 'positive' for poor welfare was then estimated.

In test selection, there are practical considerations regarding sensitivity as well as specificity. Firstly, and most importantly, we want to identify as many farms with poor welfare (according to the gold standard) as possible. But, secondly, we also want to avoid false positive results, in order to spare resources (for example, if the consequence is that herds that test positive will receive a visit by advisory services). Consequently, in our selection of a suitable set of 'tests', we applied a systematic procedure where we first ranked our tests by sensitivity within each of the three cut-off levels. We sequentially evaluated each potential welfare indicator with respect to case herds identified and, starting with the welfare indicator with the highest sensitivity, the smallest set of indicators able to pin-point all (or the largest number of) poor welfare herds was identified. In the following step, all three sets of 'tests' (the 20–80th percentile cut-off set, the 10–90th set and the 5–95th set) were jointly considered and ranked by specificity. For potential welfare indicators that, after the first selection (based on sensitivity), appeared in more than one cut-off set, the one with the highest specificity was retained. Subsequently, and in similar fashion to the procedure used in the previous step, the subset that maximised the number of herds identified while minimising the number of false positives was identified. This was called Set 1. As mortality has been suggested as an important animal-based measurement to include when assessing welfare (Winckler 2003; Fraser 2004), mortality of cows and young stock (with 95th percentile cut-off levels) were subsequently added to form Sets 2 and 3, respectively. Also, as mortality is one of the few animal-based measurements available on farms not affiliated to the Official Milk Recording Scheme, a fourth set, including only mortalities in different age categories (calves 0–24 h, calves 2–6 months, young stock 6–15 months and cows), was created.

By this non-statistical but systematic selection procedure, we identified four sets of potential welfare indicators that were then applied to the study herds, and a herd was regarded as being a case if it was positive on one or more of the tests. The overall performance of each specific set of welfare indicators was evaluated in terms of sensitivity, specificity and percentage of all herds classified correctly with respect to welfare status (poor vs not poor).

Results

Distribution of welfare remarks forming the gold standard

Descriptive statistics for the animal-based measurements are given in Table 3. The range between 90th and 100th percentiles constituted, on average, 38% (23–55%) of the total range in herd prevalence of the ten animal-based measurements. Of the 55 herds in the study population, 13 met the criteria for being classified as a herd with poor welfare. Of the remaining 42 herds, 19 had one remark. Of these, five had a remark for avoidance distance. In fact, herds with a remark on avoidance distance turned out to form a special group in that none had a remark for any of the other animal-based welfare measures. This phenomenon was not observed for any of the other animal-based measurements. For this reason, and due also to subsequent findings that cows showing a high avoidance distance were significantly more prevalent in loose-housing systems, which made up less than one third of the farms in this study, it was decided to exclude it from the set of animal-based measurements that constituted our gold standard. The distribution of number of welfare remarks, after omitting avoidance distance, is shown in Figure 1 and was 13 herds with two or more remarks, 14 herds with one remark and 28 with no remark.

Associations between animal-based welfare measurements and potential welfare indicators

Twenty-eight of the initial 65 potential welfare indicators showed a significant univariable association ($P < 0.05$) with one or more of the nine animal-based welfare measurements (Tables 4a and b). They were distributed within all focus areas except for drug use. Following variable reduction using multivariable regression, another 10 indicators were excluded leaving 18 potential welfare indicators for the systematic selection procedure (Table 4a).

Performance of welfare indicators in identifying herds at risk of having poor welfare

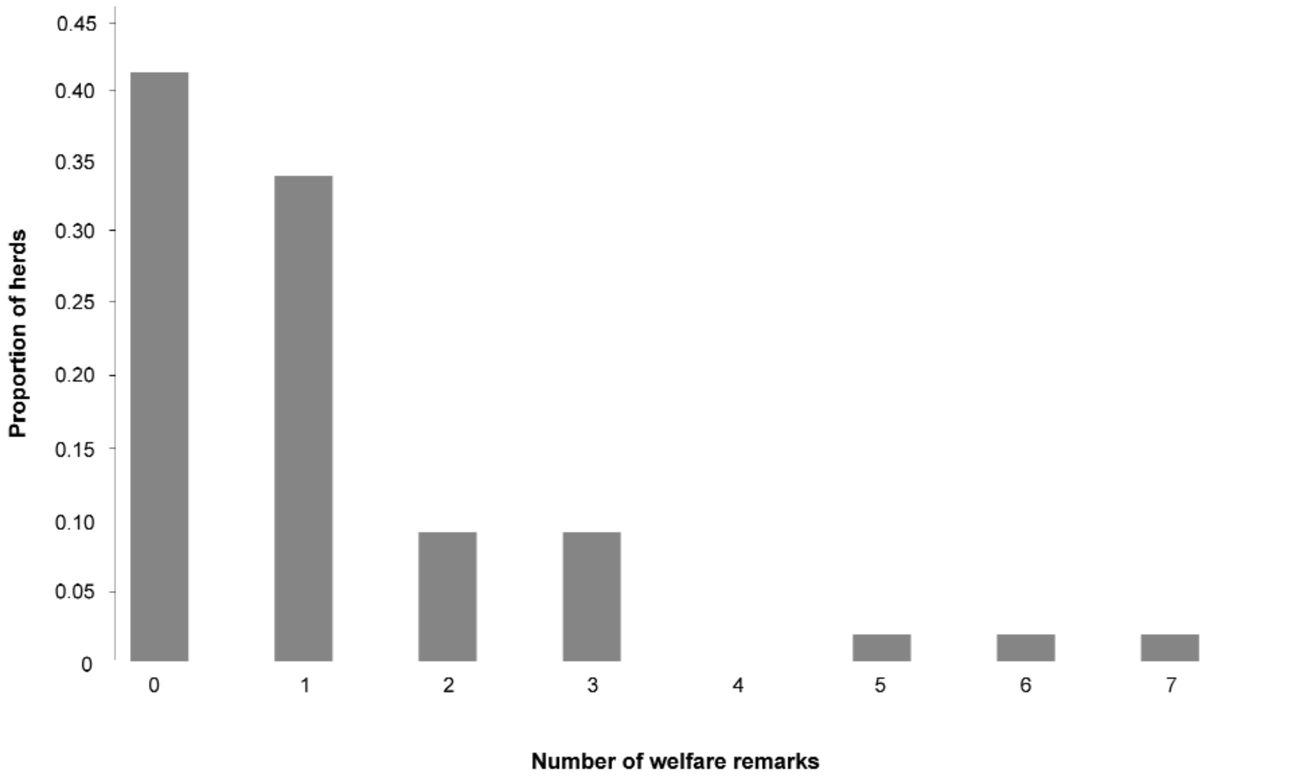
The individual sensitivities and specificities of the 18 dichotomised potential welfare indicators, at each of three cut-off levels, are shown in Table 5. After the systematic selection procedure, three welfare indicators that were jointly able to identify 8 of the 13 case herds were identified. These were: percentage cows with late ongoing artificial inseminations (> 120 days) (with a cut-off at the 95th percentile); percentage heifers without mating/artificial insemination by 17 months of age (cut-off, 95th percentile) and calf mortality 2–6 months (cut-off, 90th percentile). Descriptive statistics for the welfare indicators used to form the four sets of indicators are given in Table 6.

Table 7 shows the proportion of different welfare remarks detected by each of the four different sets of potential welfare indicators. The four sets identified on average 61, 67, 73 and 57% of the remarks, respectively. Table 8 provides the estimates of the overall performance of the different sets in terms of the percentage of herds correctly classified, sensitivity and specificity. The proportion of herds testing positive

Table 3 Herd prevalence of welfare-related conditions based on measurements on individual animals in different age groups. The recording was performed as a part of a study on welfare in Swedish dairy herds (n = 55), performed in 2005.

Herd prevalence	Minimum	10th percentile	50th percentile	90th percentile	Maximum
Cows with injuries and inflammations	0	3.5	13.4	29.2	42.6
Lame cows	0	0	2.4	6.2	10.2
Lean cows	0	0	3.0	13.2	30.4
Dirty cows	0	5.5	26.2	57.6	85.9
Cows with rising difficulties	0	0	1.4	5.9	20.3
Cows with high avoidance distance	0	1.4	6.5	23.8	62.0
Lean young stock	0	0.8	10.9	28.6	48.6
Dirty young stock	0	7.0	40.3	75.0	96.9
Lean calves	0	4.4	20.2	40.0	74.6
Dirty calves	0	1.7	25.2	47.9	69.6

Figure 1



The distribution of number of welfare remarks on nine animal-based measurements[1] used to form a gold standard for defining poor welfare[2] in a study involving 55 Swedish dairy herds. The animal-based measurements were; cleanliness and body condition in calves, cows and young stock, as well as lameness, injuries/inflammations and rising behaviour (in cows only). [1] Score above the 90th percentile = among the 10% worst. [2] A herd with ≥ 2 measurements above the 90th percentile was considered to be a herd with poor welfare.

and the predictive value of a positive test are also shown. Sets 2 and 3 detected 69 and 77% of the herds with poor welfare, respectively. All four sets correctly classified the three herds that had more than four remarks among the nine animal-based welfare measurements. Due to their slightly higher sensitivity, set numbers 2 and 3 were considered more interesting to work further with than sets 1 and 4. Set number 3 identified 20 (36%) of the herds as being at risk for having poor welfare. Seventy-three percent of all animal-based measurement remarks were found in these 20 herds.

Herds designated as having poor welfare by the gold standard, but falsely classified as not having a welfare deficiency (n = 5) had between 2 and 3 remarks on the animal-based measurements. Of the herds falsely classified as cases

https://doi.org/10.1017/S0962728600000944 Published online by Cambridge University Press

Table 4(a) Direction of significant associations between potential welfare indicators and nine animal-based welfare measurements (expressed as herd prevalences) used to form a gold standard for defining a welfare deficiency in a study involving 55 Swedish dairy herds.

Focus area	Potential welfare indicator	Dirty cows	Lean cows	Lameness	Injuries/ inflammation	Impaired rising	Lean calves	Lean young stock	Dirty calves	Dirty young stock
Management	Average calving interval, days		+[1]		(–)[2]					
	CV[3] calving interval		–[4]		(–)					
	Cows with late ongoing AIs, > 120 days			+				+		+
Calves and young stock	Calf mortality, 0–24 h					+				
	Calf mortality, 1–90 days	(+)				+				
	Calf mortality, 2–6 months		(+)				+	(+)		
	CV average milk per day (kg), 2–3rd month, primiparous cows			+						+
	Survival early lactating primiparous cows					–				
	Heifers not bred > 17 months		(+)				(+)	+	+	
Intensity of production	Cow mortality							+		
	Cows ≥ 2nd lactation									+
	Voluntary culling								+	
	Survival, early in lactating cows							+		
Feeding	Incidence of other feed related diseases including gastrointestinal disturbances	+					+	(+)		
	Prevalence of cows with urea remarks		+		(–)					
Udder health	Incidence of mastitis treatment		+							
	Prevalence of cows with chronically high cell counts								+	
Claw and leg health	Culling for claw and leg disorders				+		(–)		(–)	

[1] No brackets indicate a multivariable association ($P < 0.05$) with the animal-based measurement.
[2] Brackets indicate a univariable association ($P < 0.05$) with the animal-based measurement.
[3] Coefficient of variation.
[4] Opposite univariable direction.

(n = 8) five actually had one remark. This suggests that wrongly classified herds may have been borderline cases.

Discussion

The results of this study show that approximately 75% of the welfare remarks based on the nine animal-based measurements were found in 24% of the herds (n = 13). That is to say, a quarter of the herds accounted for three-quarters of the welfare problems. Out of 65 potential welfare indicators identified from the data register, three jointly identified 8 of the 13 case herds and 61% of the welfare remarks. These were: percentage cows with late ongoing artificial inseminations; percentage heifers without mating/artificial insemination by 17 months of age, and calf mortality in the age group 2–6 months. The results support the view that parameters from the Swedish national dairy databases, when selected and combined appropriately, can serve as a cost-effective diagnostic tool for detecting herds with poor animal welfare.

Table 4(b) Potential welfare indicators with at least one univariable, but no multivariable, significant ($P < 0.05$) association with nine animal-based welfare parameters (expressed as herd prevalences) used to form a gold standard for defining a welfare deficiency in a study involving 55 Swedish dairy herds. Direction of significant association indicated by +/–.

Potential welfare indicator	Dirty cows	Lean cows	Lameness	Injuries/ inflammation	Impaired rising	Lean calves	Lean young stock	Dirty calves	Dirty young stock
Cows with late start of AI, > 70 days									(–)
Live born calves per AI series						(–)			
Average no inseminations per heifer							(–)		
No remarks at slaughter				(–)					
Voluntary culling								(+)	
Culling overall					(+)				
Prevalence of cows with low urea levels						(–)		(–)	
Prevalence of cows with high urea levels		(+)					(+)		
Incidence risk, udder infections		(+)							
Bulk milk SCC , delivered milk		(+)						(+)	

Table 5 Point estimates for sensitivities (Se) and specificities (Sp) of potential welfare indicators, in relation to their individual ability to classify herds at risk of having a welfare deficiency in a study involving 55 Swedish dairy herds.

Welfare indicator	Cut off 95/5[1]		Cut off 90/10		Cut off 80/20	
	Se	Sp	Se	Sp	Se	Sp
Average calving interval, days	0.08	0.98	0.08	0.90	0.31	0.88
CV[2] calving interval	0.00	0.93	0.08	0.88	0.15	0.76
Cows with late ongoing AIs, > 120 days	0.23*	1.00**	0.23*	0.93	0.46*	0.79
Calf mortality, 0–24 h	0.00	0.93	0.15*	0.90	0.23	0.79
Calf mortality, 1–90 days	0.08	0.98	0.08	0.88	0.15	0.79
Calf mortality, 2–6 months	0.15*	0.98	0.31*	0.95**	0.38	0.83
Heifers not bred > 17 months	0.15*	0.98**	0.23*	0.93	0.38	0.86
CV[2] kg milk, 2-3 month primiparous cows	0.08*	0.95	0.15*	0.90	0.23	0.81
Survival early lactation, primiparous cows	0.15	0.98	0.23	0.93	0.46*	0.83
Voluntary culling	0.08	0.95	0.15	0.90	0.23	0.79
Cow mortality	0.15	0.95	0.23	0.90	0.46*	0.83
Cows ≥ 2nd lactation	0.15	0.95	0.23	0.90	0.24	0.79
Survival in early lactating cows	0.31*	0.90	0.31*	0.90	0.38	0.86
Incidence of other feed-related diseases including gastrointestinal disturbances	0.08	0.95	0.23	0.90	0.46*	0.86
Prevalence of cows with urea remarks	0.00	0.93	0.08	0.88	0.23	0.79
Incidence of mastitis treatment	0.08	0.95	0.08	0.86	0.23	0.79
Prevalence of cows with chronically high cell counts	0.08	0.95	0.08	0.88	0.31	0.76
Culling for claw and leg disorders	0.15*	0.93	0.15*	0.88	0.15	0.79

The parameters were evaluated at three different cut-offs with respect to percentile of their distributions.
* Indicates, for Se, the smallest set of potential welfare indicators that within each cut-off was able to identify most case herds.
** Indicates the final set of register parameters that was able to identify 8 of 13 case herds.
[1] 95/5, 90/10 and 80/20 indicate cut-off percentile.
[2] Coefficient of variation.

https://doi.org/10.1017/S0962728600000944 Published online by Cambridge University Press

Table 6 Descriptive statistics for potential welfare indicators, used in four different combinations as a tool to identify herds with poor welfare. The herds (n = 55) participated in a Swedish study on welfare in dairy herds, performed in 2005.

Potential welfare indicator	Minimum	80th	90th	95th	Maximum
Cows with late ongoing AIs, > 120 days[1,2,3]	0	11	13	14*	22
Heifers not bred >17 months [1,2,3]	0	75	92	100*	100
Calf mortality, 2–6 months[1,2,3,4]	0	2	3*	6	13
Cow mortality[2,3,4]	0	10	15*†	16*	25
Young stock mortality, 6–15 months[3,4]	0	2	6	7*	20
Calf mortality, 0–24 h[4]	0	9	11	13*	15

Welfare indicators were dichotomised and evaluated at three different cut-offs by percentile (80th, 90th, 95th).
* Indicates the cut-off used when indicators were combined into sets.
† In set 4, a 90th percentile cut-off was used.
A common superscript denotes indicators used in the same set.

Table 7 Percentage of welfare remarks identified by four test sets based on potential welfare indicators retrieved from pre-collected register data. In a Swedish study on welfare in dairy herds.

Animal-based measurement	Set 1[1]	Set 2[2]	Set 3[3]	Set 4[4]
Dirty cows	67	67	83	67
Lean cows	71	71	71	57
Lame cows	50	50	67	50
Cows with injuries and inflammations	50	67	67	50
Cows with rising difficulties	71	71	71	43
Lean calves	67	67	67	67
Lean young stock	67	83	83	83
Dirty calves	67	67	67	50
Dirty young stock	43	57	71	43

[1] Test set 1: cows with late ongoing AIs, > 120 days; calf mortality, 2–6 months; heifers not bred > 17 months.
[2] Test set 2: cows with late ongoing AIs, > 120 days; calf mortality, 2–6 months; heifers not bred > 17 months; cow mortality.
[3] Test set 3: cows with late ongoing AIs, > 120 days; calf mortality, 2–6 months; heifers not bred > 17 months; cow mortality; young stock mortality, 6–15 months.
[4] Test set 4: cow mortality; young stock mortality, 6–15 months; calf mortality, 2–6 months; calf mortality, 0–24 h.

Table 8 Test performance of four different sets of welfare indicators, used as test tools to identify herds at risk of having poor welfare. Cut-offs were applied to the distributions of the welfare indicators to produce a 0/1 test result, and these were combined in different sets that were identified through a systematic selection procedure. The parameter sets were applied to 55 Swedish dairy herds involved in a study on dairy cow welfare. The gold standard consisted of 9 animal-based measurements, where a herd with 2 or more values above the 90th percentile were regarded as having a welfare deficiency.

Performance parameter	Set1[1]	Set 2[2]	Set 3[3]	Set 4[4]
Correctly classified	77	76	76	77
Sensitivity	0.62	0.69	0.77	0.62
Specificity	0.87	0.89	0.91	0.87
Test positive (%)	29	33	36	29
Predictive value positive	0.5	0.5	0.5	0.5

[1] Test set 1: cows with late ongoing AIs, > 120 days; calf mortality, 2–6 months; heifers not bred > 17 months.
[2] Test set 2: cows with late ongoing AIs, > 120 days; calf mortality, 2–6 months; heifers not bred > 17 months; cow mortality.
[3] Test set 3: cows with late ongoing AIs, > 120 days; calf mortality, 2–6 months; heifers not bred > 17 months; cow mortality; young stock mortality, 6–15 months.
[4] Test set 4: cow mortality; young stock mortality, 6–15 months; calf mortality, 2–6 months; calf mortality, 0–24 h.

https://doi.org/10.1017/S0962728600000944 Published online by Cambridge University Press

Our definition of 'poor welfare' was derived by applying cut-offs based on the distribution of the herd-level prevalences of the animal-based measurements. A score among the highest 10% gave the herd a 'remark' for that measure and a herd with more than one remark was considered to be of poor welfare. Consequently, this approach characterises our definition of 'poor welfare' as being relative to other herds in the target population (primarily Swedish dairy herds and possibly dairy herds in other countries with similar management conditions). The principle of using population distributions to define targets for improvement is, however, applicable within any country and system.

In this study, we found indicators reflecting fertility problems and high mortality in different age categories to be potential indicators of poor welfare in a herd. Although we agree with objections to a direct causal relationship between poor fertility and/or high mortality rates (Rushen *et al* 2007) and the overall welfare status of a dairy herd, we interpret our results to suggest that fertility and early mortality data provide valuable information regarding welfare by their broad ability to reflect stockmanship and management in the dairy herd. In line with this, previous studies indicate a relationship between stress and fertility and also an association between a preceding welfare problem, such as lameness, milk fever, mastitis, calving difficulties, metritis or retained placenta, and reduced fertility (Dobson & Smith 2000; Dobson *et al* 2001). More specifically, we would like to suggest that consistently high mortality rates and/or poor fertility may be an indication of failure by the stockperson in monitoring and/or acting on signals of animal performance and that this general failure has a wide range of negative consequences for the welfare status of the animals on that farm. Indeed, recent cow mortality studies indicate associations at the herd level between cow and calf mortality and also between cow mortality and a limited set of welfare measurements (Thomsen *et al* 2006, 2007).

With regard to methodology, one might argue that a wider range of animal-based measures should have been used than the ten, ultimately nine, chosen in this study and wish to discuss whether or not it was appropriate to later omit avoidance distance because it did not contribute to the total number of case herds. One might also argue that the selection of potential welfare indicators could have been different or the selection procedure modified. These are, of course, valid concerns that should be addressed in future studies in this area and some will be reconsidered following experiences implementing the Scheme. In this initial attempt to use a national dairy database to identify herds with poor welfare, emphasis was placed on animal-based measures that have a broad acceptance amongst researchers and farmers, and on reliably reported database parameters that are generally agreed to be at least indirectly linked to poor animal health and welfare.

Another objection might be that a more comprehensive approach could have been to create some sort of integrated welfare index based on the nine animal-based measures. We do not agree that this would have been the appropriate approach in this case for two reasons. Firstly, in terms of acceptance by stakeholders (farmers), the stepwise approach chosen is far more transparent and, in practice, much more useful for the farmer and the advisory services. However, for our own interest, we did in fact produce a compound welfare index where the nine parameters were included, which produced similar results but where the approach chosen resulted in the selection of a slightly less efficient set of 'tests'. Secondly, welfare predictors will often be associated with specific aspects of welfare — these relationships may be lost in an index. Finally, one could also question why we chose not to use logistic regression modelling (with the gold standard as the outcome) to estimate the effect of significant welfare indicators on the odds of being a case herd according to our definition of poor welfare, and the reason is the same — this would not produce the transparent type of classification tool desired by the stakeholders. As stated earlier, the aim of this study was not to investigate the cause and effect relationships between pre-recorded register data and welfare measures, but to identify a tool that can be used to target advisory services where they might most be needed. That is to say, to farms at risk of having poor welfare.

In summary, it is clear that repeating the study using an updated list of animal-based welfare measures to determine the gold standard and a new, larger batch of farms would both add to the reliability of this approach. Associations will also change over time as animals, housing and management change, so any set of potential welfare indicators should be checked in any case at regular intervals and cut-off levels adjusted accordingly. With these cautionary notes, we would like to suggest that an approach such as the one described in this paper might be applied beneficially in other countries. The type of register data in this study is probably feasible to collect in many countries around the world and may already be available in a number of areas. Since associations will probably differ between countries for a number of reasons, similar models will have to be developed in the environment where they are to be used and any set of potential welfare indicators will of course also need to be validated regularly against on-farm, animal-based measures.

Animal welfare implications

The results of this study imply that farms at risk of having poor welfare might be successfully identified by using pre-recorded data. This would result in a more effective targeting of advisory services by channelling visits to those farms that may be in most need of such advice. Depending on the model used, up to 77% of farms could be identified correctly (according to the gold standard based upon visits to the actual farms) with a sensitivity of up to 0.77. Depending on what pre-recorded data is available in other countries it may be possible to repeat the approach taken here to improve the process by which herds at risk of poor welfare are identified.

References

Bartussek H 1999 A review of the animal needs index (ANI) for the assessment of animals' well-being in the housing systems for Austrian proprietary products and legislation. *Livestock Production Science 61*: 179-192

Brolund L 1985 Cell counts in bovine milk. Causes of variation and applicability for diagnosis of subclinical mastitis. *Acta Veterinaria Scandinavica Supplement 80*: 1-123

Chaplin S and Munksgaard L 2001 Evaluation of a simple method for assessment of rising behaviour in tethered cows. *Animal Science 72*: 191-197

Dobson H and Smith RF 2000 What is stress and how does it affect reproduction? *Animal Reproduction Science 60-61*: 743-752

Dobson H, Tebble JE, Smith RF and Ward W 2001 Is stress really that important? *Theriogenology 55*: 65-73

Emanuelson U 1988 The national Swedish animal disease recording system. *Acta Veterinaria Scandinavica Supplement 84*: 262-264

Fraser D 2004 Applying science to animal welfare standards. *Global Conference on Animal Welfare: an OIE Initiative* pp 121-134. 23-25 February 2004

Gillund P, Reksen O, Karlberg K, Randby, ÅT, Engeland I and Lutnæs B 1999 Utprøving av en holdvurderingsmetode på NRF-kyr. *Norsk Veterinærtidskrift 111*: 623-632. [Title translation: Evaluation of a method for body condition scoring of Norwegian Red Dairy Cows]

Gustafsson AH 1993 *Acetone and urea concentration in milk as indicators of the nutritional status and the composition of the diet of dairy cows*. Dissertation, Swedish University of Agricultural Sciences, Department of Animal Nutrition and Management, Uppsala, Sweden

Keeling LJ and Veissier I 2005 Developing a monitoring system to assess welfare quality in cattle pigs and chickens. *Science and Society Improving Animal Welfare*. *Welfare Quality Conference Proceedings* pp 46-50. 17-18 Nov 2005, Brussels, Belgium

Olsson SO, Baekbo P, Hansson SO, Rautala H and Osteras O 2001 Disease recording systems and herd health schemes for production diseases. *Acta Veterinaria Scandinavica 42(1)*: 51-60

Philipsson J, Eriksson JÅ and Stålhammar H 2003 Know-how transfer in animal breeding, the power of integrated cow data bases for farmer´s selection of bulls to improve functional traits in dairy cows. *EAAP Cattle Commission Session V*. 31 Aug-3 Sept 2003, Rome, Italy

Rousing T 2003 *Welfare assessment in dairy cattle herds with loose-housing cubicle systems Development and evaluation of welfare indicators*. PhD Thesis, The Royal Veterinary and Agricultural University, Copenhagen and Dept of Animal Health and Welfare, Copenhagen, Denmark

Rousing T and Waiblinger S 2004 Evaluation of on-farm methods for testing the human-animal relationship in dairy herds with cubicle loose housing systems, test-retest and inter-observer reliability and consistency to familiarity of test person *Applied Animal Behaviour Science 85*: 215-231

Rushen J, de Passillé AM, Keyserlingk MAG and Weary D 2007 *The Welfare of Cattle*. Springer: Dordrecht, The Netherlands

Sprecher DJ, Hostetler DE and Kaneene JB 1997 A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology 47*: 1179-1187

Swedish Animal Welfare Act 1988 *Djurskyddslagen: L1, (Animal Welfare Act: L1) SFS 1988: 534*. www/sjv.se, Swedish board of agriculture

Thomsen PT, Kjeldsen AM, Sørensen JT, Houe H and Ersbøll AK 2006 Herd-level risk factors for the mortality of cows in Danish dairy herds *Veterinary Record 158*: 622-626

Thomsen PT, Østergaard S, Houe H and Sørensen JT 2007 Loser cows in Danish dairy herds: Risk factors *Preventive Veterinary Medicine 79*: 136-154

Veissier I and Evans A 2007 Rationale behind the welfare quality assessment of animal welfare. *Assuring Animal Welfare: From Societal Concerns to Implementation, Second Welfare Quality Conference Stakeholder Conference* pp 9-12. 2-3 May 2002, Berlin, Germany

Webster AJF and Main DCJ 2003 Proceedings of the 2nd International Workshop on the Assessment of Animal Welfare at Farm and Group Level. *Animal Welfare 12(4)*

Winckler C, Baumgartner J and Waiblinger S 2007 Proceedings of the 3rd International Workshop on the Assessment of Animal Welfare at Farm and Group Level. *Animal Welfare 16(2)*

Winckler C, Capdeville J, Gebresenbet G, Hørning B, Roiha U, Tosi M and Waiblinger S 2003 Selection of parameters for on-farm welfare assessment protocols in cattle and buffalo. *Animal Welfare 12*: 619-624

Whay HR, Main DCJ, Green LE and Webster AJF 2003 Assessment of the welfare of dairy cattle using animal-based measurements: Direct observations and investigation of farm records. *The Veterinary Record 153*: 197–202

https://doi.org/10.1017/S0962728600000944 Published online by Cambridge University Press