The Old School, Brewhouse Hill, Wheathampstead,
Hertfordshire AL4 8AN, UK
www.ufaw.org.uk

Animal Welfare 2019, 28: 307-315
ISSN 0962-7286
doi: 10.7120/09627286.28.3.307

The representativeness of a semi-random sampling method for animal welfare assessments on mink farms

AF Marsbøll, BIF Henriksen and SH Møller*

Department of Animal Science, Aarhus University, Blichers Allé 20, PO Box 50, DK-8830 Tjele, Denmark
* Contact for correspondence: steenh.moller@anis.au.dk

Abstract

*In this study we present a semi-random sampling method developed for the sampling of mink (*Neovison vison*) for on-farm welfare assessments according to the WelFur-Mink system. The only information required for implementation of this method is the number of cages in use in each shed on the farm. The representativeness of samples selected with this method was evaluated in relation to the physical characteristics of the farm and the mink characteristics by simulated sampling on a farm with a complicated structure in the growth period. The selection of 10,000 samples was simulated. The trueness was, in general, high, ie the method has no systematic skewness. The precision was low for certain factors due to the high variation within sheds. The sampling in sections of six adjacent cages means that it is often not possible to select a sample which is an exact representation of the mink and their housing environment. If accepting a deviation of ± one cage section, the estimated probability of selecting a representative sample was high for most of the individual factors. However, the estimated probability of selecting a sample that is representative according to all factors was rather low. This deviation from exact representativeness ought to be evaluated in the light of the increased feasibility and repeatability offered by the method. Also, we expect that the representativeness of samples selected with this method will be higher on other less-complicated farms. We suggest that this simple method balances feasibility and representative sampling in a way that makes it useful in the WelFur-Mink system.*

Keywords: *animal welfare, feasibility, mink, on-farm welfare assessment, simulation study, WelFur*

Introduction

When assessing animal welfare at farm level, it is usually not possible to include all the animals on the farm since it would be too time consuming and, thus, unfeasible for practical reasons. Hence, welfare assessments are often based on a sample of the animals, and the welfare of the animals in the sample is considered to be an estimate of the welfare of the animals on the farm. It is, therefore, important that the sample is representative, ie that the animals and housing conditions of the animals in the sample reflect those on the farm. Different animal welfare assessment systems sample according to different rules depending on species and production system. In some systems, a random selection of the animals is suggested as, for example, random selection from microchip numbers in horses (*Equus caballus*) (AWIN 2015). However, when selecting a number of sows (*Sus scrofa domesticus*) in a large pen, the sows cannot easily be identified. Therefore, a more practical approach was developed in Welfare Quality®: the first sow in sight is the starting sow. The next sow is "the sow whose head is the fourth away (facing) from the 'starting sow'", and so forth (Welfare Quality® 2009a). Many sampling strategies also include some kind of stratification. This means that the animals in the sample should be distributed so that the relevant subgroups on the farm are represented in the sample according to the size of the sub-group, which often increases the representativeness of the sample compared to random sampling (Lohr 2010). One example is dairy cows (*Bos taurus*) kept in different groups where the number of cows sampled in each group should be proportionate to the size of the group (Welfare Quality® 2009b).

In the on-farm welfare assessment system, WelFur for foxes (*Vulpes vulpes*, *V lagopus*) and mink (*Neovison vison*), the assessment is based on a sample of the animals on the farm. This system is largely inspired by the Welfare Quality® project (Mononen *et al* 2012). WelFur was implemented in all European mink and fox farms on a voluntary basis starting from January 2017, and the resulting assessments are used for welfare certification of the pelts from these farms (Fur Europe 2017). As in Welfare Quality®, animal-based measurements are preferred over resource- and management-based measurements (Mononen *et al* 2012), and in the assessment protocol for mink (WelFur-Mink) nine out of 22 measurements are animal-based (Møller *et al* 2015). Mink are seasonal breeders, and the strict annual production system can be divided into three seasons, each associated with specific animal groups: i) winter season with adult breeders on the farms; ii) nursing season with mostly adult females and kits on the farms; and iii) growth season with mostly adult females and juveniles on the farms (Møller *et al* 2003). Due to the seasonal production system, the full welfare assessment according to the WelFur-Mink protocol includes one assessment in each season. In each season, a sample of mink is selected and used for the assessment of all animal- and resource-based measurements (Møller *et al* 2015). Mink farms typically consist of several sheds in which the mink are kept in rows of cages elevated off the ground. The cages are generally constructed as battery cages with approximately 2 m between the bearing posts. This means that each row can be physically divided into battery cage sections which, depending on the design, often consist of 5 to 8 cages each (Jørgensen 1985). In order to increase the feasibility of the assessment, and especially the assessment of stereotypic behaviour where 2 min of observation is required, the sample is selected in sections of six adjacent cages. The sample consists of 15 sections of six adjacent cages (90 cages) in the growth period and 20 sections of six adjacent cages (120 cages) in the winter and nursing period (Møller *et al* 2015). Each cage section in the sample thus represents 6.7% of the sample in the growth period (six out of 90 cages) and 5% in the winter and nursing period (six out of 120 cages). In the winter period, there are only adult breeders on the farms. They are usually kept individually, hence the sample in this period consists of 120 mink. In the nursing and growth period, the sample consists of a varying number of mink. In the nursing period, most cages hold one adult female and her litter, as most of the males and unmated females are pelted after the mating period. In the growth period, adult females are typically housed individually or with one or two juvenile kits, while the rest of the juveniles are typically housed in pairs or groups.

In WelFur-Mink, the original sampling method used stratification in order to ensure a representative sample. The subgrouping factors were sex, age and colour type in the nursing and winter period and sex, age, social housing conditions and colour type in the growth period. Besides this, different types of sheds, cages, nest-boxes and watering systems should also be considered (Møller *et al* 2015). Due to the many subgrouping factors in the stratification, a lot of information was needed beforehand. Practical tests of WelFur-Mink before implementation showed that retrieving this information could be a challenge. In case the farmers did not have the information, the feasibility of the assessment would be impaired if the assessors should attempt to collect the information themselves, as this would be impossible to do within one day on most farms. In case it was possible to get the information needed, the practical test also showed that it was difficult and time-consuming to select the cage sections on the farm in a way that made the resulting sample representative according to all factors. Hence, the original sampling method was not applicable in practice. Furthermore, the individual assessor was responsible for selecting the cage sections on the farm based on the stratification. Thus, the resulting samples could be unintentionally biased. In order to overcome these problems we developed a new sampling method. This new method takes the systematic structure of mink farms into account and the only information that is needed is the number of cages in use in each shed on the farm. The sampling of the cage sections in the sample is based on: i) a systematic distribution between the sheds according to the number of cages in use in each shed; and ii) a random selection within the sheds. This method thus combines interval sampling (which may also be referred to as systematic sampling) and random sampling (Lohr 2010) and is, therefore, not fully random; hence, it can be considered a semi-random sampling method. A preliminary version of the method was described by Marsbøll and colleagues (2016). In the present paper, we present an improved and revised method.

The purpose of this study was to describe this new semi-random sampling method and to explore the representativeness of samples taken with this method as well as the probability of taking a representative sample. This was done by simulated sampling in the growth period on a farm with a complicated structure. We choose the growth period due to the social housing of the mink and because almost all cages are in use during this period, which makes sampling more challenging than in the other periods.

Materials and methods

Model farm

A model farm was used to test the sampling method. This model farm mimicked the set-up of a private Danish mink farm, thus ensuring a realistic model, and was also chosen due to its complexity. This complexity was made up both by the size of the farm as well as the many different combinations of social housing, colour types and housing conditions on the farm. Such complexity would make it particularly challenging to use the original sampling procedure described in the WelFur-Mink assessment protocol. Mink are farmed under similar conditions across Europe, but farm size and complexity differ significantly within and between countries. The model farm may be considered a worst-case scenario for sampling on mink farms in Europe. In October 2015, each battery cage section on the private farm was described according to the minks' characteristics, the social housing conditions and the housing environment. If there were any differences within the battery cage sections, the most frequent situation was noted. This description of the private farm formed the basis for the model farm.

On this model farm, there were approximately 42,000 cages divided between 23 sheds and about 98,000 mink. As typical for this period, there were no adult males on the farm. The ratio between the adult females and juveniles was 1:5 and there were several different colour types. The number of cages in use in each shed ranged from 500 to 6,800. In some sheds, there were only mink in one stage (eg adult females) or one colour type. In others, there were mink at several stages (eg adult females and juveniles) and colour types housed in

Table 1 An example of the calculations that form the basis of the semi-random sampling method for on-farm welfare assessments according to the WelFur-Mink system.

Total number of sheds: 7		Number of cage sections in sample: 15	
Total number of cages in use: 3,900		**Sample threshold[†] = 3,900/15 = 260**	
Shed order[‡]	**Cages in use**	**Ratio[§]**	**Cage sections to select[#]**
5	400	400/260 = 1.54	2–0 = 2
6	400	(400+400)/260 = 3.08	3–2 = 1
7	1,500	(400+400+1,500)/260 = 8.85	9–3 = 6
1	400	(400+400+1,500+400)/260 = 10.38	10–9 = 1
2	400	(400+400+1,500+400+400)/260 = 11.92	12–10 = 2
3	400	(400+400+1,500+400+400+400)/260 = 13.46	13–12 = 1
4	400	(400+400+1,500+400+400+400+400)/260 = 15.00	15–13 = 2

[†] The total number of cages in use on the farm divided by the number of cage sections in the sample;
[‡] A randomly picked shed number identifies which shed the starting one in the calculations and the rest of the sheds follow in sequential order;
[§] The number of cages in use is summarised for the step-wise addition of each shed and divided by the sample threshold;
[#] The ratio is rounded to its nearest integer and the number of cage sections that have been selected in the previous sheds is subtracted.

different combinations (eg one adult female with one juvenile male, or one juvenile female with one juvenile male). Housing conditions differed between the sheds, and some housing-related factors (eg cage type, nest-box insulation and nest-box position) also differed within individual sheds.

Sampling method

The only information needed for using the new sampling method on a mink farm is the number of cages in use in each shed on the farm. We expect the farmers can provide this information but, if not, the assessors can collect it themselves. This will still be time-consuming, but possible within the limitations of one workday for each welfare assessment. As an example, the sampling of 15 cage sections in the growth period on a farm with mink in seven sheds is shown in Table 1.

First, each shed is given a sequential number. Most mink farms have a systematic layout and the numbering of the sheds should follow these systematics. For example, if there are two-row sheds in one area and multi-row sheds in another, all the sheds in one area should be numbered before the sheds in the other area. Next, a randomly picked shed number identifies which shed is the starting one in the calculations. The sample threshold is the total number of cages in use on the farm divided by the number of cage sections (eg 15 or 20) in the sample. The ratio between the sample threshold and the number of cages in use is calculated for the step-wise addition of each shed. This means that first the ratio between the sample threshold and the number of cages in use in the starting shed is calculated. This ratio is rounded to its nearest integer, and this number is the number of cage sections to be randomly selected within the starting shed. Subsequently, the ratio between the sample threshold and the number of cages in use in the starting shed plus the following shed is calculated. This ratio is rounded to its nearest integer, and this number minus the number of cage sections that has been selected in previous sheds, is the number of cage sections to be randomly selected in the second shed. This procedure is repeated until the number of cages in use in all sheds has been added. The random selection in the respective sheds is based on the sheds' physical division into battery cage sections of 5 to 8 cages. If the calculations show that one cage section should be selected in a shed, which can be divided into, eg 50 battery cage sections, one of the 50 battery cage sections is randomly selected and included in the sample. However, as the sample is based on sections of six adjacent cages, if the battery sections consist of, eg five cages, an additional cage must be included from one of the neighbouring physical sections, and if the battery section consists of, eg seven cages, one cage must be omitted.

Testing the sampling method

A programme that was able to simulate the new sampling method was developed in R (R Core Team 2017). The selection of 10,000 samples on the model farm was simulated in order to have a sufficient number of samples to examine the representativeness of samples selected with the method. Each sample consisted of 15 sections of six adjacent cages as in the growth period in WelFur-Mink, and the selection of each individual sample was independent of previously selected samples. The prevalence of mink in relation to the levels of the sub-grouping factors included in the original stratification in WelFur-Mink was calculated for the simulated samples and for the model farm (farm prevalence). The sub-grouping factors and their levels are listed in Table 2. Only factors and levels that showed variability within the model farm were included in the analysis. This means that the following factors and levels were omitted:

Table 2 The sub-grouping factors and their levels included in the original stratification in WelFur-Mink (adjusted from Møller *et al* 2015).

Factor	Level
Sex and age	Adult males
	Adult females
	Juveniles
Housing of adult dams	Individually
	One or two juvenile males
	With their litter
	With other adults
Housing of juveniles	Individually
	Male-female pairs
	Other groups
Colour types	Brown
	Mahogany
	Black
	Cross
	Palomino
	Pearl
	Silver blue
	White
	Other/mixed
Cage type	Single/pair
	Group
Cage wall	Solid
	Wire mesh
	Other
Nest-box presence	Yes
	No
Nest-box position	Normal
	Top
Nest-box insulation	High
	Medium
	Low
Watering system	Automatic with frost protection
	Automatic without frost protection
	Manual
Shed type	Two row
	Multi row

the factors 'Watering system' (all mink had access to an automatic and frost protected water supply) and 'Nest-box presence' (all mink had access to a nest-box), and the factor levels 'Adult males' (there were no adult males), 'Adult females family housed with their litter' (this housing was not used), 'Adult females housed with other adults' (this housing was not used), 'Black' (there were no mink with this colour type), 'Mahogany' (there were no mink with this colour type) and 'Cage wall other than wire mesh or solid' (all cage walls were wire mesh or solid).

The representativeness of the method was evaluated based on trueness and precision. Trueness refers to the closeness of agreement between the simulated samples and the actual farm value, while precision refers to the closeness of agreement between the simulated samples. Trueness and precision were assessed individually for each of the included factor levels using measures that take into account that data were not normally distributed in all cases. The trueness was assessed based on the Error, in order to evaluate whether there is a systematic bias. The Error for each factor level was calculated as the median of the sample prevalences minus the farm prevalence. The precision was assessed based on the 95% central range in order to evaluate the range of the sample distribution. The 95% central range for each factor level was calculated as the 97.5 percentile of the sample prevalences minus the 2.5 percentile of the sample prevalences.

The probability of selecting a representative sample was assessed by estimating the probability of selecting an individual sample where the prevalence of mink in the sample is within a range of ± 5, 10, 15, 20 and 25 percentage points of the farm prevalence. This was done by calculating the share of samples where the prevalence of mink in each sample is within a range of ± 5, 10, 15, 20 and 25 percentage points of the farm prevalence for each factor level and for groups of factors and their levels. All calculations were made in R (R Core Team 2017).

Results

Each sample consisted of 15 sections of six adjacent cages, ie each sample always consisted of 90 cages. The number of mink in the samples varied from 168 to 246 with a median of 210. All samples included one or more cage sections from the six largest sheds, while it varied how many cage sections the samples included from the smaller sheds. Cage sections were selected in all sheds.

The prevalence of mink on the model farm and the distribution of the prevalences in the simulated samples for each factor level are shown in Table 3. The Error was low for most of the included factors, ranging from –2.6 to 1.2 percentage points. The lowest Error was for the prevalence of the factor levels 'Adult females' and 'Juveniles', while the highest was for the prevalence of the colour type 'Palomino.' The range of the sample distributions varied from a 95% central range of 4.2 percentage points for the prevalence of the factor level 'Housing of juveniles — individually' to a 95% central range of 43.0 percentage points for the prevalence of the factor 'Housing of adults.'

https://doi.org/10.7120/109627286.28.3.307 Published online by Cambridge University Press

Table 3 The prevalence of mink on the model farm and the distribution of the prevalences in the 10, 000 simulated samples for the factor levels related to sex, age, social housing, colour type and housing conditions.

Factor	Level	Farm prevalence (%)	Error† (percentage points)	95% range‡ (percentage points)
Sex and age	Adult males	18.9	0.0	19.7
	Juveniles	81.1	0.0	19.7
Housing of adults	Individually	13.1	1.2	43.0
	One or two juvenile males	86.9	–1.2	43.0
Housing of juveniles	Individually	0.6	–0.6	4.2
	Male-female pairs	47.0	–0.6	33.9
	Other groups	52.4	0.7	33.7
Colour types	Brown	36.6	–0.5	37.9
	Cross	1.6	–1.6	11.8
	Palomino	8.5	–2.6	21.2
	Pearl	10.5	0.3	27.8
	Silver blue	15.4	–0.2	26.0
	White	17.0	0.1	22.9
	Other/mixed	10.4	–1.6	27.3
Cage type	Single/pair	54.6	–0.3	27.7
	Group	45.4	0.3	27.7
Cage wall	Solid	49.1	–0.6	15.6
	Wire mesh	50.9	0.6	15.6
Nest-box position	Normal	98.0	–0.7	9.1
	Top	2.0	0.7	9.1
Nest-box insulation	High	60.5	0.1	27.5
	Medium	38.9	–0.2	27.7
	Low	0.6	–0.6	5.6
Shed type	Two row	11.2	0.6	13.7
	Multi row	88.8	–0.6	13.7

† The difference between the farm prevalence and the median of sample prevalences;
‡ The difference between the 2.5 percentile and the 97.5 percentile of the sample prevalences.

The estimated probability of selecting a sample where the prevalence of mink in the sample is within ± 5, 10, 15, 20 or 25 percentage points of the farm prevalence for each factor level is shown in Table 4. The estimated probabilities increased with the range of the deviation from the farm prevalence. For some factor levels, the estimated probability was low at a deviation range of ± 5 percentage points but increased to a probability above 0.90 already at a ± 10 percentage points deviation range (eg the 'Palomino' colour type with an increase from an estimated probability of 0.57 to 0.94). A few factor levels had a smaller increase and did not reach an estimated probability above 0.90 until the deviation range was increased to ± 25 percentage points (eg 'Housing of adult females — individually' with an increase from an estimated probability of 0.32 at deviation range of ± 5 percentage points to 0.95 at a deviation range of ± 25 percentage points). For a few factor levels, the estimated probability was above 0.90 for a deviation range of ± 5 percentage points (ie the factors 'Nest-box position' and 'Shed type' and the factor levels 'Housing of juveniles — individually' and 'Nest-box insulation — low').

Table 4 The estimated probability of selecting a sample where the prevalence of mink are within ± 5, 10, 15, 20 or 25 percentage points of the farm prevalence for the factor levels related to sex, age, social housing, colour types and housing conditions.

Factor	Level	± 5	± 10	± 15	± 20	± 25
Sex and age	Adult females	0.68	0.95	1.00	1.00	1.00
	Juveniles	0.68	0.95	1.00	1.00	1.00
Housing of adults	Individually	0.32	0.42	0.82	0.88	0.95
	One or two juvenile males	0.32	0.42	0.82	0.88	0.95
Housing of juveniles	Individually	1.00	1.00	1.00	1.00	1.00
	Male-female pairs	0.42	0.74	0.92	0.98	1.00
	Other groups	0.44	0.74	0.92	0.98	1.00
Colour types	Brown	0.37	0.67	0.88	0.97	0.99
	Cross	0.83	0.97	0.99	1.00	1.00
	Palomino	0.57	0.94	0.99	1.00	1.00
	Pearl	0.55	0.74	0.96	0.99	1.00
	Silver blue	0.51	0.87	0.95	1.00	1.0
	White	0.57	0.89	0.99	1.00	1.00
	Other/mixed	0.45	0.74	0.96	0.99	1.00
Cage type	Single/pair	0.49	0.86	0.96	1.00	1.00
	Group	0.49	0.86	0.96	1.00	1.00
Cage wall	Solid	0.80	0.98	0.99	1.00	1.00
	Wire mesh	0.80	0.98	0.99	1.00	1.00
Nest-box position	Normal	0.95	0.99	1.00	1.00	1.00
	Top	0.95	0.99	1.00	1.00	1.00
Nest-box insulation	High	0.46	0.80	0.96	1.00	1.00
	Medium	0.47	0.82	0.96	1.00	1.00
	Low	0.98	1.00	1.00	1.00	1.00
Shed type	Two row	0.96	0.97	1.00	1.00	1.00
	Multi row	0.96	0.97	1.00	1.00	1.00

Table 5 The estimated probability of selecting a sample where the prevalence of mink are within ± 5, 10, 15, 20 or 25 percentage points of the farm prevalence for all factors in a group.

Group of factors	± 5	± 10	± 15	± 20	± 25
Sex, age and social housing	0.11	0.28	0.74	0.86	0.94
Colour types	0.02	0.29	0.75	0.93	0.99
Housing environment	0.16	0.65	0.92	0.99	1.00
All factors	0.00	0.07	0.53	0.80	0.93

https://doi.org/10.7120/109627286.28.3.307 Published online by Cambridge University Press

The estimated probability of selecting a sample where the prevalence of mink in the sample is within ± 5, 10, 15, 20 or 25 percentage points of the farm prevalence for groups of factors is shown in Table 5. The estimated probability of selecting a sample where the prevalence of mink in the sample are within a deviation range of ± 5 percentage points in all factor levels was very low, but increased to 0.52 at a deviation range of ± 15 percentage points and reached an estimated probability above 0.9 at a deviation range of ± 0.25 percentage points. The estimated probability of selecting a sample where the prevalence of mink in the sample are within a given range for all factor levels in a group, were above 0.9 at a deviation range of ± 15 percentage points for the factors related to housing environment. For the factor levels related to colour types, the probability was above 0.9 at a deviation range of ± 20 percentage points, and for the factor levels related to sex, age and housing conditions, the probability was above 0.9 at a deviation range of ± 25 percentage points.

Discussion

The systematics of the method

This study presents a new semi-random sampling method which was developed for the sampling of mink for on-farm welfare assessments according to WelFur-Mink. In WelFur-Mink, samples consisting of 15 sections of six adjacent cages in the growth period and 20 sections of six adjacent cages in the winter and nursing period are used for the on-farm assessment of welfare (Møller *et al* 2015). With our new semi-random sampling method, the cage sections in the sample are selected based on a systematic distribution between the sheds according to the number of cages in use in each shed followed by a random selection within the respective sheds. As a consequence of this method, cage section is always selected in sheds or a group of sheds where the total number of cages in use exceeds the sample threshold, ie larger 'groups' of mink and their housing environment will always be included in the sample, while the inclusion of smaller groups may vary. For example, if a farmer keeps mink with a special colour type in one part of the farm, these will always be represented in the sample if the number of mink in this group is larger than the sample threshold. The individual samples selected on the model farm presented in this study always included one or more cage sections from the largest sheds, while the inclusion of cage sections from smaller sheds varied. As more mink were housed in the larger sheds, they were also more representative for the model farm. This systematic approach, therefore, increases the representativeness of the method. Also, the risk of selecting all cage sections in the same shed or groups of sheds, as could have been the case with a completely random sampling, is excluded.

Trueness and precision

Trueness and precision were used to evaluate how well the method performed when selecting several samples on the same model farm mimicking a complicated farm in the growth period. The Error was low for most of the included factors. This means that the method has no systematic skewness, ie the systematics of the method did not increase nor reduce the probability of selecting a sample with specific characteristics. However, the range of the sample distribution varied between the factors. The farm that we used as a model had a large variation within the sheds in colour type, sex, age and the social housing of the mink but, surprisingly, also in housing conditions as there could be, eg several different cage and nest-box designs within the same shed. This has decreased the precision of the method as the random selection, in combination with a large variation, results in several different combinations of the different factors and their levels. We, therefore, expect that samples selected on farms with less variation within sheds will have a greater precision than in this study.

The probability of selecting a representative sample

In practice, only one sample is selected on each farm. The sample is selected in sections consisting of six adjacent cages where the housing conditions are similar and the mink most often share the same characteristics. As each cage section represents 6.7% of the total sample in the growth period and 5% in the winter and nursing period, it is only possible in rare cases to select a sample that is an exact representation of the mink and their housing environment. For example, if the actual prevalence of males on a farm in the winter period is 17.5%, the closest representation in the WelFur sample would be 15% (three cage sections with males) or 20% (four cage sections with males). And, as the number of mink per cage varies in the nursing and growth periods, some cage sections will also represent more than 5 and 6.7% of the total sample in these periods. Thus, no matter which sampling method is used, we cannot expect the samples to be an exact representation of the farm they are selected from. But how large a deviation from the actual farm value can be considered representative? If no deviation is accepted, the range of the closest representation will be 5 percentage points in the winter period and minimum 5 and 6.7 percentage points in the nursing and growth period, respectively. If a deviation of ± one cage section is accepted, the range of an acceptable representation is increased to 15 (± 7.5) percentage points in the winter period and minimum 15 (± 7.5) and 20.1 (± 10.1) percentage points in the nursing and growth periods, respectively. If a deviation of ± 10 percentage points is considered acceptable in the growth period, the estimated probability of selecting an acceptable representative sample in this study was 0.95 for age and sex, and between 0.80 and 1.00 for all the housing-related factors, which we consider quite a high probability. For the factor levels related to colour type, the probability ranged from 0.67 to 0.97, and for the factor levels related to social housing of females and juveniles, the probability ranged from 0.42 to 1.00. This means that the probability is quite low for some factors. However, if we increase the acceptable deviation to ± 15 percentage points, the probability of selecting a representative sample is increased, ie for the factor levels related to colour type the estimated probability ranged from 0.88 to 0.99, and for the factor levels related to social housing of females and juveniles the estimated probability ranged from 0.82 to 1.00. Thus, for individual factors, the estimated probability of selecting a representative sample is quite high, if ± 15 percentage points is considered an acceptable deviation. However, when looking at the estimated probability of selecting a sample, where the prevalence of mink in the sample are within a deviation range of ± 15 percentage points for all factor levels, the estimated probability is only 0.53. When considering groups of factors, the estimated probability of selecting a sample where the prevalence of mink in the sample are within a deviation range of ± 15 percentage points is 0.92 for the factor levels related to housing environment, 0.75 for colour types and 0.74 for sex, age and social housing. This indicates that there is a high probability that the individual sample is representative according to the factor levels related to housing environment, while the probability of a representative sample in regards to age, sex, social housing and colour type is lower. Thus, only few samples can be expected to be representative according to all factor levels, but due to the high probability of selecting a sample that is representative in regards to the individual factor levels, it can be expected that any sample as a minimum is representative according to some or most factor levels.

Choice of method

Much effort has been put into the development of valid and reliable welfare indicators, as this is crucial for a correct assessment of animal welfare at farm level (eg Mononen *et al* 2012; Veissier *et al* 2013). However, if the entire farm is not included in the assessment, a correct assessment requires that the sample is representative in relation both to the physical characteristics of the farm and the animals on it. If the sample does not represent the farm, the assessment may be misleading, even though the sample is assessed correctly using both valid and reliable welfare indicators. Also, sampling must be reliable, hence assessor bias must be avoided. In the present study, a feasible procedure for taking a sample of mink at farm level without the risk of assessor bias is presented. To our knowledge, this is the first published study evaluating a sampling method for taking a sample of animals for welfare assessments at farm level by investigating how well the samples represent the farm with regards to the characteristics of the animals and their housing environment. Other studies have evaluated sampling strategies by assessing how the welfare of the animals in the sample reflects the welfare at farm level (eg Main *et al* 2010). In our case, this was not possible as it would have taken us several months to assess the welfare of the 98,000 mink on the model farm. On a smaller farm, this would have been possible, but we aimed to test the method in a worst-case scenario, ie with as many different combinations of housing systems, type of animals, etc as possible. We, therefore, believe that evaluating the performance of the methods as regards the welfare of the mink in the sample should be based on simulating different prevalences of welfare problems on a model farm.

Using the semi-random sampling method in WelFur-Mink

The samples selected in this study were not all acceptable representations of the farm, even if a deviation of ± one cage section was accepted. However, this deviation from exact representativeness ought to be evaluated in light of the increased feasibility and reduced risk of assessor bias, which the method offers. Also, according to our experience from welfare assessments across Europe, the farm that was used as a model in the study was more complicated (ie more variation within and between sheds) than most farms. The farm was chosen for that exact reason, and this study is, therefore, a 'worst-case scenario.' Hence, we expect that the representativeness of samples selected with this method will be higher on other less-complicated farms. Based on the results and discussions above we, therefore, suggest this semi-random sampling method as a reliable way to balance feasibility and representative sampling in WelFur-Mink.

Animal welfare implications

On-farm welfare assessment can be used to determine animal welfare status of farms, eg for their certification. The resulting assessments can also be used for advising farmers on what to do in order to enhance animal welfare on their farm (Main *et al* 2003). However, all procedures must be highly feasible for welfare assessments to be useful in practice, and the time constraint is often a challenge as shown, for example, in dairy cattle where the estimated time needed for a Welfare Quality® assessment in a herd with 200 cows is 7.7 h (Knierim & Winckler 2009). From an economic point of view, it is also important to reduce the time needed for a full assessment (Sørensen *et al* 2007). The implementation of WelFur-Mink in all European mink farms started in January 2017. The results of the welfare assessments are to be used for certification of farms and to provide decision support for the farmers. If all farms sign up for the assessment, the welfare of mink on approximately 3,000 mink farms will be assessed annually in at least one of the three seasons. This feasible sampling method contributed to making this large-scale implementation possible, with the potential to improve the welfare of mink in Europe. The method can also be used for selecting representative samples of mink in other countries and for other purposes as, for example, health inspection by veterinarians or official inspection of compliance with rules and regulation by authorities. It may also be useful for sampling in other animal production systems where animals within distinct units are more similar than between units.

Acknowledgements

We thank the owners of the private farm for kindly allowing us to use their farm as a model and Tarja Koistinen, Eeva Ojala and Elna Mortensen for testing the first version of sampling method in practice. All authors were responsible for development and revision of the method and contributed to the manuscript. AFM was responsible for data collection, data management, simulated sampling and generating the results. Fur Europe and a PhD grant from Aarhus University, Denmark financed this project.

References

AWIN 2015 *AWIN welfare assessment protocol for horses*. 10.13130/AWIN_horses_2015

Fur Europe 2017 *4,000 European fur farms to be WelFur certified by 2020*. http://www.fureurope.eu/news/4-000-european-fur-farms-to-be-welfur-certified-by-2020/

Jørgensen G 1985 *Mink Production*. Scientifur: Tjele, Denmark

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

Lohr LS 2010 *Sampling: Design and Analysis, Second Edition*. Brooks/Cole: Boston, USA

Main DCJ, Barker ZE, Leach KA, Bell NJ, Whay HR and Browne WJ 2010 Sampling strategies for monitoring lameness in dairy cattle. *Journal of Dairy Science 93*: 1970-1978. https://doi.org 10.3168/jds.2009-2500

Main DCJ, Kent JP, Wemelsfelder F, Ofner E and Tuyttens FAM 2003 Applications for methods of on farm welfare assessment. *Animal Welfare 12*: 523-528

Marsbøll AF, Henriksen B and Møller SH 2016 It is possible to take a representative sample of animals based on the number of cages in use in each mink shed. In: Mäki-Tanila P, Valaja J, Mononen J, Sironen T and Vapalahti P (eds) *Proceedings of the XIth International Scientific Congress in Fur Animal production* pp 337-342. 23-26 August 2016, Helsinki, Finland

Møller SH, Hansen SW, Malmkvist J, Vinke CM, Lidfors L, Gaborit M and Botreau R 2015 *WelFur Welfare assessment protocol for mink*. Fur Europe: Brussels, Belgium

Møller SH, Hansen SW and Sørensen JT 2003 Assessing animal welfare in a strictly synchronous production system: The Mink Case. *Animal Welfare 12*: 699-703

Mononen J, Møller SH, Hansen SW, Hovland AL, Koistinen T, Lidfors L, Malmkvist J, Vinke CM and Ahola L 2012 The development of on-farm welfare assessment protocols for foxes and mink: the WelFur project. *Animal Welfare 21*: 363-371. https://doi.org/10.7120/09627286.21.3.363

R Core Team 2017 *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria. https://www.R-project.org/

Sørensen JT, Rousing T, Møller SH, Bonde M and Hegelund L 2007 On-farm welfare assessment systems: what are the recording costs? *Animal Welfare 16*: 237-239

Veissier I, Winckler C, Velarde A, Butterworth A, Dalmau A and Keeling L 2013 Development of welfare measures and protocols for the collection of data on farms or at slaughter. In: Blokhuis H, Miele M, Veissier I and Jones B (eds) *Improving Farm Animal Welfare: Science and Society Working Together: The Welfare Quality Approach* pp 115-146. Wageningen Academic Publishers: Wageningen, The Netherlands. https://doi.org/10.3920/978-90-8686-770-7_6

Welfare Quality® 2009a *Welfare Quality® assessment protocol for pigs*. Welfare Quality® Consortium: Lelystad, The Netherlands

Welfare Quality® 2009b *Welfare Quality® assessment protocol for cattle*. Welfare Quality® Consortium: Lelystad, The Netherlands