European Food Safety Authority

EFSA Journal 2013;11(7):3299

SCIENTIFIC REPORT OF EFSA

Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses[1]

European Food Safety Authority[2, 3]

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

Council Directive 2007/43/EC[4] on the protection of chickens kept for meat production requires Member States to submit to the Commission the results of a data collection of welfare indicators in broiler slaughterhouses. The Commission prepared a draft document for a harmonised data collection system of welfare indicators in broiler slaughterhouses for a one-year period and requested EFSA to assess whether the sample size chosen in the document would be sufficient to achieve the objectives of the data collection. The structure of the population as well as possible stratification factors influence the sampling protocol of a data collection. In this case, the sampling should occur in each Member State at three levels: i) at slaughterhouses; ii) at flocks within each sampled slaughterhouse and iii) at broilers and feet within each sampled flock. Based on the methodological approach developed for this mandate, the assessment shows that for 15 MS, 1% of the total number of flocks slaughtered in one year is not a sufficient sample to assess 1% prevalence of the given conditions. For 3 MS, 1% sample is sufficient depending on the combination number of flocks and broilers/feet chosen. For 1 MS, when any combination is used, 1% of flocks slaughtered is a sufficient sample to assess prevalence of 1%. A general approach for designing and ensuring representativeness of sampling plans in MS was proposed. The recommended sampling scheme better fitting for this data collection is a three stage sampling scheme, where the selection of the unit at each level could be based on established criteria. Finally, in the case an EU harmonised data collection is implemented in the future, it is recommended to submit the collected data in a harmonised way through the EFSA Data Collection Framework (DCF). For this purpose, this report also proposes a possible data model for the submission of the data collected.

Key words

Welfare of broilers, data collection, broiler slaughterhouses, welfare indicators, sample size

---

[1] On request from the European Commission, Question No EFSA-Q-2012-01002, approved on 28 June 2013.

[2] Correspondence: ahaw@efsa.europa.eu

[3] Acknowledgement: EFSA wishes to thank the members of the Working Group on Data collection at broiler slaughterhouses: Charlotte Berg, Gerard van Breukelen, Toby Knowles and Virginie Michel for the preparatory work on this scientific output and the hearing expert: Laure Bignon and EFSA staff: Maria Ferrara, José Cortinas Abrahantes, Jane Richardson, Ana Afonso and Andrea Gervelmeyer for the support provided to this scientific output.

[4] Council Directive 2007/43/EC of 28 June 2007 laying down minimum rules for the protection of chickens kept for meat production. OJ L182, 19-28.

Suggested citation: European Food Safety Authority, 2013. Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses. EFSA Journal 2013; 11(7):3299, 68 pp. doi:10.2903/j.efsa.2013.3299

Available online: www.efsa.europa.eu/efsajournal

SUMMARY

Art 6(2) of Council Directive 2007/43/EC laying down minimum rules for the protection of chickens kept for meat production requires Member States (MS) to submit to the Commission the results of the data collection based on the monitoring of a representative sample of flocks slaughtered during a minimum period of one year. The sampling and data collected, as referred to in Annex III of the Directive, should be scientifically based, objective and allow for a comparative analysis. The data to be collected at slaughterhouse is on welfare indicators, for the assessment of the welfare on farm of broilers raised in that MS according to Council Directive 2007/43/EC.

The Commission has prepared a *draft* document proposing a harmonised data collection system of welfare indicators in broiler slaughterhouses. The main objective is to achieve an overall picture in the EU and in each of the 27 MS of the welfare of broilers, in particular detecting those welfare conditions, such as contact dermatitis, systematic illness and mortality, which occur at 1% or greater at EU level. The Commission draft document for the data collection foresees to stratify the sample in two categories: MS and maximum stocking densities. The defined sample unit is the flock and the sample size is 1% of randomly selected flocks out of the total number of flocks slaughtered each year, with a minimum of 200 flocks and a maximum of 500 flocks (in the case where 1% of the flocks represent more than 500 flocks). As the objective of the data collection is to assess prevalence of given welfare conditions at the MS level, the flocks to be sampled should be farmed in the given MS, according to Council Directive 2007/43/EC.

EFSA was requested to assess if the proposed sample size is sufficient to achieve the objectives set.

To calculate the sample size needed for this data collection exercise, the prevalence, the variability parameters and the test characteristics chosen need to be defined. The EFSA methodological scientific report "Sample Size Considerations for Hierarchical Populations" (EFSA, 2013) published in connection with this scientific report, outlines the methodological approach that was developed for this mandate.

The sampling of broilers should occur in each MS at three levels: first, slaughterhouses; second, flocks within each sampled slaughterhouse and third, broilers and feet within each sampled flock. Moreover, as the sampling is carried out at more than one level, between-slaughterhouses and between-flocks variabilities need to be taken into account.

In the case of a complex hierarchical population, as described above, analytical expressions do not exist to estimate sample sizes needed. Therefore, theoretical approximations as well as computer simulation exercises were carried out to estimate the sufficient sample sizes by MS. The theoretical derivation highlighted that when a MS has less than 20 slaughterhouses, all of them should be sampled, applying a two level model where the sampling refers only to the number of flocks sampled per slaughterhouse and to the number of broilers/feet sampled per flock. When in a MS the sample size is smaller than 10% of the total number of flocks slaughtered, the sufficient sample size to detect 1% marginal prevalence is 360 flocks slaughtered in the country in one-year period, with 100 broilers/feet sampled for each flock. When in a MS with less than 20 slaughterhouses the sample size is higher than 10% of the total number of flocks slaughtered in the country and a finite population correction is needed, the calculation was carried out using the three level theoretical formula specifically developed for this mandate.

In order to reply to this technical assistance mandate, EFSA requested to MS the necessary data for the assessment. Data at country level to perform the sample size assessment was provided by 19 MS; 15 MS provided full data also at slaughterhouse level, allowing the preparation of specific sampling plans. The parameters chosen were i) prevalence to be detected of 1% ii) effect size of 0.5% iii) power of the test of 80% with 95% level of confidence iv) variability at slaughter level of 1 and at flock level of 2.5 for the three level model, and v) variability at flock level of 4.2 for the two level

model. Although the chosen variability parameters were considered appropriate in order to follow a precautionary approach, it should be noted that different values could be used for the calculation.

The results of the assessment show that for 15 MS, 1% of the total number of flocks slaughtered in one-year period is not a sufficient sample to assess 1% prevalence of the given conditions. For 3 MS, the 1% sample is sufficient depending on the combination number of flocks and broilers/feet chosen. For 1 MS, when any combination is used, 1% of the total number of flocks slaughtered in one-year time is a sufficient sample to assess a prevalence of 1%. For those MS where 1% of the total number of flocks slaughtered in one year was found to be an insufficient sample, EFSA has provided alternative samples sizes.

Depending on various defined scenarios of prevalence and prevalence variations between slaughterhouses and flocks, EFSA has developed a tailored three level formula to allow the calculation of the number of units to be sampled for one level, given defined sample sizes at each of the other two levels.

When conducting a survey, the method used to calculate the sample size to be tested depends on the purpose of the survey, the design prevalence of the condition to be examined, the level of confidence of the survey results and the performance of the test used. The structure of the population under consideration influences the sampling protocol and the modelling of the data.

EFSA has proposed a general approach for designing sampling plans in MS for this data collection, in order to ensure representativeness of the sample chosen. The recommended sampling scheme better fitting for this one-year exercise has been considered by EFSA as a three stage sampling scheme, where the selection of the unit at each level could be based on the following criteria:

i) at stage I (slaughterhouse level), the selection at this step depends on the number of slaughterhouses processing broilers in the MS;

ii) at stage II (flock level), the selection at this stage is proposed to be a fixed number whenever the number of slaughterhouses is larger than 20, in other cases a proportion of the flocks processed per slaughterhouse is recommended to be sampled;

iii) at the stage III (broilers/feet level), the number of broilers or feet to be sampled is fixed for all flocks sampled within a MS.

On the basis of the above-described approach, specific sampling plans have been provided for 14 MS where complete data at slaughterhouse level were available.

Finally, in the case that an EU harmonised data collection is implemented in the future and MS submit data directly to EFSA, it is recommended to submit such data in a harmonised way through the EFSA Data Collection Framework (DCF). For this purpose, this report also proposes a possible data model for the submission of the data collected at slaughterhouse through the DCF.

TABLE OF CONTENTS

Suggested citation: European Food Safety Authority, 2013. Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses. EFSA Journal 2013; 11(7):3299, 68 pp. doi:10.2903/j.efsa.2013.3299

Available online: www.efsa.europa.eu/efsajournal

BACKGROUND AS PROVIDED BY THE EUROPEAN COMMISSION

Council Directive 2007/43/EC[5] on the protection of chickens kept for meat production provides in Article 6(2) that Member States shall submit to the Commission the results of data collection of welfare indicators based on monitoring of a representative sample of broilers' flocks slaughtered during a period of one year.

In order to help Member States to implement Article 6(2) of the Directive and to have harmonised data on welfare indicators of broilers, the Commission has prepared a *draft* document proposing a harmonised data collection system of welfare indicators in broiler slaughterhouses. It was elaborated on the basis of EFSA opinions on the welfare of broilers[6] and of a meeting organised by the Commission with scientific experts. The document was thoroughly discussed with Member States representatives during two dedicated working groups attended by EFSA representatives.

The *draft* document proposing a harmonised data collection system of welfare indicators in broiler slaughterhouses is attached.

TERMS OF REFERENCE AS PROVIDED BY THE EUROPEAN COMMISSION

The Commission would like to request EFSA, in the framework of Article 31 of Regulation (EC) No 178/2002 to assist the Commission by indicating whether a percentage of 1% of the total number of flocks slaughtered each year in each Member State, with a minimum of 200 flocks would ensure that the proposed data collection system of welfare indicators will comply with the Council Directive 2007/43/EC.

In case such a sample would not meet the objectives of Article 6(2) of the Directive, the Commission requests EFSA to provide possible alternatives of a representative sample of broiler flocks slaughtered during a period of 12 months in each Member State and at EU level. The Commission highlights the importance to strive for simplification of the data collection system and to avoid the introduction of burdensome additional requirements.

The Commission requests that the EFSA technical assistance should be made available by April 2013. EFSA will present possible choices of representative sample of flocks slaughtered during a period of 12 months in each Member State and at EU level to the Standing Committee of the Food Chain and Animal Health of April/May 2013. The Commission may use the results, together with comments from the Member States for the drafting of a Commission Implementing Decision on data collection in broiler slaughterhouses.

[5] COUNCIL DIRECTIVE 2007/43/EC of 28 June 2007 laying down minimum rules for the protection of chickens kept for meat production. OJ L182, 19-28.
[6] EFSA Journal 2010; 8 (7):1666 [82 pp.]; EFSA Journal 2010; 8(7):1667 [81 pp.].

CONTEXT OF THE SCIENTIFIC OUTPUT

Art 6(2) of Council Directive 2007/43/EC laying down minimum rules for the protection of chickens kept for meat production requires Member States (MS) to submit to the Commission the results of the data collection based on the monitoring of a representative sample of flocks slaughtered during a minimum period of one year.

The sampling and data collected, as referred to in Annex III of the Directive, should be scientifically based, objective and allow a comparable analysis. The data to be collected at slaughterhouse are welfare indicators, which should allow assessing the welfare on the farm of broilers raised according to Council Directive 2007/43/EC. Such monitoring during post-mortem inspection would allow feeding back information from the slaughterhouse to the farm, to take the necessary corrective measures.

On the basis of national experiences in particular from 6 MS (Sweden, Denmark, United Kingdom, France, Spain and Portugal) which had already put in place pilot data collection systems, the Commission has elaborated a draft document proposing a harmonised data collection system of welfare indicators in EU broiler slaughterhouses.

The Commission draft document is enclosed in Appendix A of this report. The Commission draft document proposes a data collection designed as a one-year period exercise, in which only criteria of on farm welfare are monitored, with the main objective of achieving an overall picture in the EU and in each of the 27 MS of broiler welfare, in particular for abnormal levels of conditions such as contact dermatitis, systematic illness and mortality. It is foreseen to stratify the sample in two categories: MS and maximum stocking densities. The actual stocking density of the flock at the moment of departure to slaughterhouse is to be recorded. Data should be collected in at least 5 different slaughterhouses; in MS where there are less than 5 slaughterhouses, all of them should be the sampled. Seasonality is also to be taken into account with data to be collected at least once a month and analysed throughout the year. With the objective of detecting the prevalence of conditions which occur at 1% or greater at EU level, the sample unit is the flock and the sample size is 1% of randomly selected flocks out of the total number of flocks slaughtered each year, with a minimum of 200 flocks and a maximum of 500 flocks (in the case where 1% of the flocks represent more than 500 flocks). In MS where less than 200 flocks are slaughtered each year, all flocks should be tested. Moreover, 300 birds per flock for hock burn and 100 randomly selected feet per flock for foot pad dermatitis, are to be assessed.

The seven welfare indicators proposed in the Commission's draft document (hock burn, foot pad dermatitis, total rejections, culling rate, cumulative daily mortality, standard mortality and percentage of birds being dead-on-arrival at the slaughterhouse) were selected on the basis of their feasibility, animal welfare impact (in terms of number of birds affected and severity of the condition) and their correlation with management practices. Hock burn, foot pad dermatitis, total rejections and dead on arrival are to be evaluated and recorded directly at the slaughterhouse by the operators or the competent authorities; culling rate, cumulative daily mortality and standard mortality are recorded by the owners and provided at the slaughterhouse. Council Directive 2007/43/EC already requires for cumulative daily mortality and dead on arrival at the slaughterhouse to be recorded and provided at slaughterhouse in case of flocks with stocking densities higher than 33kg/m^2.

ANALYSIS

1. OBJECTIVES

Following discussions with the 6 MS and clarifications from the EU Commission, the approach to the mandate, as presented in this report, was based on the following points:

i) The percentage of flocks (1%) to be sampled in each MS out of the total number of flocks slaughtered in one year in the MS is to be considered, instead of 1% of flocks calculated at overall EU level;
ii) The 1% of flocks slaughtered each year in each MS should be tested with the objective to detect and assess prevalence of conditions which occur at 1% (or greater) at MS level;
iii) The above-mentioned conditions to be detected by the data collection are those represented by the welfare indicators listed in the Commission draft document (hock burn, foot pad dermatitis, total rejections, culling rate, cumulative daily mortality, standard mortality and dead on arrival at slaughterhouse);
iv) If 1% of the total number of flocks slaughtered in each MS is found to be not representative, EFSA will propose alternatives of sample sizes of slaughtered flocks for each MS, that would ensure representativeness.

The objective of this scientific report is to reply to the Commission technical assistance mandate on the representativeness of the sample sizes proposed in the Commission draft document for a data collection in broiler slaughterhouses.

In particular, this report:

i) assesses if 1% sample size for each MS is sufficient to detect 1% prevalence of the conditions indicated in the draft Commission document;

ii) proposes a tailored statistical three level formula to allow calculation, in various defined scenarios of prevalence and prevalence variations between slaughterhouses and flocks, of the number of units to be sampled at slaughterhouse, flock and broilers/feet levels;

iii) proposes alternative representative samples, as found necessary;

iv) outlines the steps to design possible sampling protocols that would ensure representativeness of the sample as well as specific sampling plans for those MS for which full data were available;

v) in the case that an EU harmonised data collection will be implemented in the future, and if data will be submitted to EFSA, it proposes a possible data model for the submission of the data collected at the slaughterhouse directly to EFSA, through the EFSA Data Collection Framework (DCF).

The EFSA methodological Scientific report "Sample Size Considerations for Hierarchical Populations" (EFSA, 2013) published in connection with this Scientific report, outlines the full methodological approach that was developed for this mandate. The methodological aspects of this report, in reply to the Commission mandate, are broadly explained in the above-mentioned EFSA methodological report and are supported with specific examples and formulas. The specific formulas will be referred to in this report, as necessary.

2. INTRODUCTION

When developing a monitoring system for animal welfare, in this case broiler welfare, an approach similar to epidemiological health monitoring could be taken into account. In general, a monitoring system is made up of the following steps: i) identification of the goal; ii) identification of the population concerned, and definition and selection of the survey population; and iii) selection of the indicators and the systematic collection of data. Moreover, in animal health research, the visualisation of links between risk factors and health outcomes is based on harmonised collection of standardised indicators, followed by analysis to investigate relevant associations. If some animal-based measures of broiler welfare were to be collected in a systematic manner, then this would pave the way to investigate not only associations among factors, consequences (welfare conditions) and animal-based measures (indicators), but also the predictive capacity of those correlations and associations (EFSA, 2012b). In this case, appropriate sample sizes and sampling protocols would need to be designed to allow a meaningful analysis of the collected data, depending on the set objectives.

The crucial factors when deciding on an indicator are that it is valid (it really says something about the welfare of the birds) and that it can be measured reliably (by different people and under different conditions). If it is going to be applied in practice, it is also necessary for the measure to be feasible (EFSA, 2010). In addition when a particular measure is taken will influence the results and consideration may be given to taking it at the most critical point in time. How exactly the measure is taken will also influence the results, and whether it is based on a sample of birds or not. If the measure is based on a sample of birds, how these are selected is important and will need to be standardized if results are to be comparable (EFSA, 2010).

Moreover, in general when conducting a survey, the method used to calculate the sample size to be tested depends on the purpose of the survey, the expected prevalence of the condition to be examined (the so called "design prevalence"), the level of confidence of the survey results and the performance of the test used (EFSA, 2012d). The structure of the population under consideration clearly influences the sampling protocol and the modelling of the data as well as possible stratification factors that could lead to design of a stratified sampling protocol.

3. THE METHODOLOGICAL APPROACH TO THE DATA COLLECTION IN BROILER SLAUGHTERHOUSES

3.1. Defining the sampling protocol best fitted for the purpose: stratified versus simple random scheme

According to the Commission draft document outlining the harmonised data collection, the sample should be stratified into two categories: MS and stocking density, implying that the broiler population should be pre-stratified by MS and also in terms of stocking density. Moreover, the sampling of broilers should occur in each MS at 3 levels: first, slaughterhouses are sampled; second, flocks within each sampled slaughterhouse are sampled; third and last, broilers and feet within each sampled flock are sampled.

In a stratified sample of the population, a random sample is taken from each non-overlapping stratum/subpopulation, which together form the entire population.

Based on Council Directive 2007/43/EC and on stocking densities applied by MS, the following categories should be considered as the strata for a possible pre-stratification of the population:

Cat 1: ≤ 33 kg/m^2
Cat 2: 33 kg/m^2 $< x \leq 36$ kg/m^2
Cat 3: 36 kg/m^2 $< x \leq 39$ kg/m^2
Cat 4: 39 kg/m^2 $< x \leq 42$ kg/m^2

Prestratification in this case would ensure that each stratum is sufficiently sampled and would easily allow a comparison of prevalence of the conditions between strata. However, this approach requires that within a given MS there is *a priori* knowledge of the stocking density of each flock before arrival at the slaughterhouse. Moreover, when the main objective of the survey is to estimate the overall prevalence of certain welfare conditions at MS level rather than a comparison among strata of the same population, then prestratification can lead to biased prevalence estimates, unless the distribution of broiler flocks by stocking density category is exactly known.

Prestratification on MS is advisable as it ensures sufficiently precise estimation of the welfare conditions identified by the indicators from the data collected. But, contrarily, prestratification on maximum stocking density would lead to a biased estimation of the indicator prevalence at MS level, unless the frequency distribution of the classes of stocking densities in the MS is precisely known as well as it is being stable between successive months.

Therefore, instead of prestratification on maximum stocking density, **a three stage random sampling protocol** for each MS is recommended for this specific data collection exercise, since it ensures that the frequency distribution of maximum stocking density in the sample reflects the distribution in the population of that MS, excluding possible sampling errors. Full details are provided in Sections 2.1.1, 2.1.2 and 2.1.3 of the EFSA Scientific methodological report (EFSA, 2013).

As the sampling is carried out at more than one level, between-slaughterhouse and between-flock variabilities need to be taken into account. Prevalence of a condition/indicator at population level can be defined as i) conditional, being the probability that any given broiler has the condition, given no specific effect at the level of flocks and at the level of slaughterhouses or ii) marginal, being the probability that any given broiler, from any given flock within any given slaughterhouse, has the condition, given certain effects at flocks and slaughterhouses level. Whilst the interest in this data collection is focused on marginal prevalence at MS level of 1%, methodological results for sample size and power calculations are based on conditional models. Therefore a transformation from marginal to conditional prevalence is necessary to calculate the sample size. The formula for this transformation is provided in Section 2.4 (equations 13 and 14) of the EFSA methodological report (EFSA, 2013). In the absence of effects at flock and slaughterhouse level, conditional and marginal prevalence are the same.

The geographical distribution of slaughterhouses in a MS as well as differences in throughput may, to some extent, lead to differences between slaughterhouses in the prevalence of the welfare conditions under investigation. This may depend on factors which include regional differences in climate, broiler genotype, housing conditions (such as ventilation type, litter materials, floor type) and feeding.

In order to generalise from the prevalence in the sampled slaughterhouses to the prevalence at MS level, slaughterhouses should be treated as a random effect, implying that the sampled slaughterhouses should be a random sample of all slaughterhouses in the MS. For this particular data collection, sampling of a small number of slaughterhouses per MS (as indicated in the draft Commission document and depending on the MS size) would allow a representative assessment of prevalence level at MS level, as long as the sampled slaughterhouses are either reasonably representative of all slaughterhouses in the MS in terms of region and capacity, or that they account for the vast majority of the annual throughput in that MS (EFSA, 2013). In the case where the sampled slaughterhouses vary significantly in terms of annual throughput of broiler flocks processed, throughput variations and possible prevalence variations between slaughterhouses need to be taken into account. In this case, in order to obtain an unbiased prevalence estimate at MS level, the most precise approach would be to sample the same percentage of the annual throughput of each slaughterhouse participating in the data collection. While this sampling scheme would lead to a more precise estimation of the prevalence of the variable of interest, an alternative, and feasible approach, would be to sample the same number of flocks and the same number of broilers/feet per flock in all slaughterhouses. The appropriate estimation method that accounts for differences between slaughterhouses and between flocks in size as

well as in prevalence should then be applied. The formula for this calculation is provided in Sections 2.1.2 and 2.1.3 of the EFSA methodological report.

In the case that the detailed information regarding slaughterhouses and their throughput, as well as flock sizes are available, other sampling schemes could be used to gain precision in the estimation process without the requirement to weight the estimates. Examples of such sampling schemes are provided for specific MS in Appendix C of this report. The general methodology is described in the EFSA methodological report (EFSA, 2013).

As flocks may differ in size (commonly 10.000 to 30.000 broilers) (EFSA, 2012a) and in prevalence of conditions, in order to obtain an unbiased prevalence estimate at the MS level, all animals sent to slaughter in that MS in the time frame of the data collection should have the same probability of being sampled. Sampling flocks with equal probabilities and then sampling the same number of broilers from each sampled flock is a practical approach, but it is known that this procedure can induce bias since broilers from larger flocks then have a smaller probability of being sampled (EFSA, 2013). However, this bias can be eliminated if an appropriate approach is taken when analysing the data, in particular by weighting the flock-specific prevalence by flock size. The formula to weight the flock-specific prevalence estimate by flock size to obtain the overall prevalence estimate is provided in Section 2.1.3 of the EFSA methodological report. As mentioned above, a similar complication arises at the slaughterhouse level if slaughterhouses differ in size and prevalence: sampling the same number of flocks per slaughterhouse irrespective of slaughterhouse size implies that flocks from larger slaughterhouses have a smaller probability of being sampled. Unbiased prevalence estimation then requires weighting the slaughterhouse-specific prevalence by slaughterhouse size in terms of the annual number of flocks slaughtered.

The recommended sampling scheme better fitting for this one-year exercise and, as mentioned above, is a **three stage sampling scheme.** The three stages are: i) at stage I (slaughterhouse level), the selection at this step will depend on the number of slaughterhouses processing broilers in the MS; ii) at stage II (flock level), the selection at this stage is proposed to be a fixed number whenever the number of slaughterhouses are larger than 20; if there are less than 20 slaughterhouses, it is recommended to sample proportions of the flocks processed per slaughterhouse; iii) at the stage III (broilers/feet level) the number of broilers or feet to be sampled is fixed for all flocks sampled within a MS. More information is provided in Section 5 and Appendix C of this report.

3.2. Assessment of the main elements of the sampling protocols

3.2.1. Sample size: the statistical formulation of the problem

It should be noted that, for the assessment of the indicators foreseen in the Commission draft document for the data collection exercise, sampling within the flock is only carried out for hock burn and foot pad dermatitis. For the recording, by flock, of dead on arrival at the slaughterhouse, total rejections, culling rate, cumulative daily mortality and standard mortality, these indicators are taken at slaughterhouses and from production records relating to the entire flock, hence no physical sampling is carried out. Moreover, the following sample size assessment, within the conditions illustrated below, shows that the same sample size of broilers/feet can be tested to assess both foot pad dermatitis and hock burn and that a different sample size is not necessary.

In the case of a complex hierarchical population, as it is for this data collection exercise, analytical expressions do not exist to directly provide sample size estimates. Therefore, theoretical approximations as well as computer simulation exercises needed to be undertaken. All the above considerations were taken into account in developing the sample size estimation procedures for complex populations, in reply to Commission's mandate.

The question from the Commission has been re-formulated in statistical terms as a null hypothesis H_0, indicating that for a given welfare condition/indicator in a MS population (i.e. y= occurrence of hock burn in a MS), its prevalence π_y should be smaller or equal to the pre-specified threshold of 0.01 (1%).

$$H_0: \pi_y \leq 0.01 \ \ vs \ \ H_A: \pi_y > 0.01$$

The statistical formulation of the mandate would therefore be to assess/calculate the number of flocks and the number of broilers/feet per flock needed to estimate the prevalence of the welfare conditions indicated in the Commission draft document with a certain level of precision, and to have sufficient power to reject the null hypothesis if the true prevalence reaches a certain level above 0.01, for instance 0.015.

In hypothesis testing, in general, the significance level is set to 5%, and the power of the test to a value of 80%. Proceeding with the statistical problem formulation, a **three level logistic model** was used to estimate the parameters needed to model the potential scenarios, that could then be used as a basis to estimate the required sample size. Datasets on prevalence of foot pad dermatitis and hock burn from specific pilot studies carried out in UK and France (Haslam et al, 2008; Allain et al, 2009; EFSA, 2013) were used to fit the model and to estimate the specific between-slaughterhouse and between-flock variabilities, which were both found to be significantly different from 0. These two variabilities are important parameters for determining the sample size.

Based on the general considerations above, and on the analysis of the data available from the pilot studies, the population structure to consider in each MS has been defined as a three level population (Slaughterhouses; Flocks and Broilers/Feet). In order to estimate the appropriate sample size, an effect size should be also defined, which in this case was considered to be 0.005 (0.5%). This means that the sample size must be sufficient to reject the null hypothesis of a prevalence smaller or equal than 0.01 (1.0%) if the true prevalence is 0.015 (1.5%). This strict effect size was chosen to take a conservative approach, although other values could be used. Although simulations were run also on a two level model (Flocks and Broilers/Feet), the analysis of the available pilot studies and the between-slaughterhouses variability found for foot pad dermatitis indicates that it is not advisable to consider only a two level hierarchy of the population, but that the slaughterhouse level should also be considered.

Specific characteristics of the population, in particular the variation of the prevalence for hock burn and foot pad dermatitis, were simulated at slaughterhouse level and at flock level using the estimates from the pilot studies mentioned above. In particular, the variability associated with the slaughterhouses was considered at $\sigma_S^2 = 1$, while for flocks two values were considered ($\sigma_F^2 = 2.5$ and $\sigma_F^2 = 4.2$), all based on the results obtained from the mixed effects logistic regression analysis of the pilot studies. Full details of the variance analysis are provided in Section 2.2 of the EFSA methodological report (EFSA, 2013). It should be noted that, although the variability can be different in other MS, the values used are following a precautionary approach. Simulations were used to generate a three level hierarchical structure based on a population prevalence of 0.015 (1.5%, which considered the threshold of 1% together with the established effect size 0.5%). Considering the marginal prevalence, the model fitted, together with the estimated prevalence $(\widehat{\pi_y})$ obtained for each simulated data, was used to assess if such estimated parameter was statistically larger than the specified threshold of 0.01 (therefore rejecting H_0).

In the simulations, the number of slaughterhouses to be sampled in each MS was considered to be 10 or 20, with the number of flocks sampled ranging from 10 to 50 (10, 20, 30, 40 and 50) within each of the sampled slaughterhouse, and with the number of broilers sampled in each flock being 100, 200, 300, 400 and 500. The variability parameters used for the simulations were $\sigma_S^2 = 1$ for slaughterhouses, and $\sigma_F^2 = 2.5$ and $\sigma_F^2 = 4.2$ for flocks as indicated above. For each combination of samples indicated above for the three levels as well as for each variability associated to

slaughterhouses and flocks, 1000 simulations were run. The power of the test was calculated based on the number of times the null hypothesis was rejected out of the total number of simulations run.

In particular it should be noted that, with the parameters considered, the sampling of only 10 slaughterhouses, even when sampling 500 broilers within 50 flocks per slaughterhouse, would not allow achieving the necessary power to detect a marginal prevalence above 0.01 (EFSA, 2013).

Figures 1 and 2 show the power curves and the number of flocks and broilers/feet to be sampled, considering the selected prevalence variability parameters and when 20 slaughterhouses are sampled, ignoring finite population correction (see section later on).

In particular, Figure 1 shows that considering a variability of 2.5 at the flock level, the number of flocks needed per slaughterhouse to reach 80% power with 95% confidence is 20, with at least 200 broilers/feet sampled per flock.

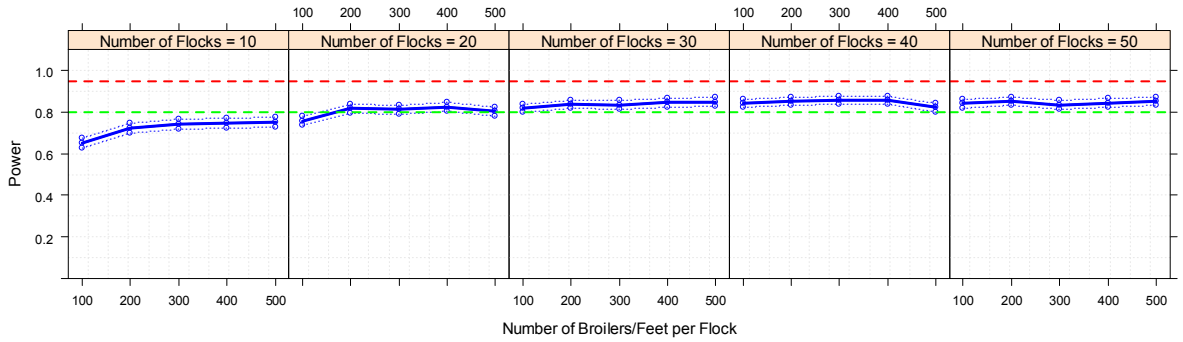


Figure 1: Power curve for each simulated scenario considering 20 slaughterhouses sampled, variability at slaughterhouse level of 1 and at flock level of 2.5, and effect size of 0.005 to detect population prevalence greater than 0.01. The green and red dash lines represent the power levels of 0.8 and 0.95 respectively.

Figure 2 shows that considering a variability of 4.2 at the flock level, the number of flocks needed per slaughterhouse to reach 80% power with 95% confidence is 20, with at least 300 broilers/feet sampled per flock.

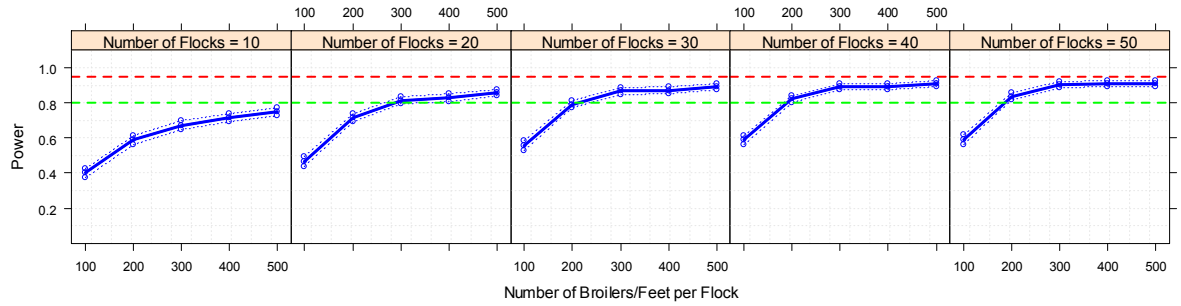


Figure 2: Power curve for each simulated scenario considering 20 slaughterhouses sampled, variability at slaughterhouse level of 1 and at flock level of 4.2, and effect size of 0.005 to detect population prevalence greater than 0.01. The green and red dash lines represent power levels of 0.8 and 0.95 respectively.

In parallel to the above-mentioned simulations, theoretical formulas to estimate the necessary numbers of slaughterhouses, flocks and broilers to be sampled in a three level hierarchical population, were also derived. The results of the theoretical derivations from the formulas were compared to the simulated scenarios (Figures 1 and 2), considering the parameters described above; the comparison indicated agreement between the computations of both approaches, after corrections due to the approximation used to derive the theoretical formulas, as explained in Section 2.3.2 of the EFSA methodological report (EFSA, 2013).

Such derived theoretical formulas, once the variability values for the main indicators (hock burn and foot pad dermatitis) at slaughterhouse and at flock levels are defined, allows calculation of:

- The number of slaughterhouses to be sampled, given the number of flocks to be sampled per slaughterhouse and the number of broilers/feet to be sampled per flock;
- The number of flocks to be sampled, given the number of slaughterhouses to be sampled and the number of broilers/feet to be sampled per flock;
- The number of broilers/feet to be sampled per flock, given the number of slaughterhouses to be sampled and the number of flocks to be sampled.

Applying this formula it is therefore possible to calculate, with a practical approach and depending on the chosen setting of the variability parameters as well as on the resources available, the number of units to be sampled for one level, given a defined sample size for the other two levels. The full theoretical formulas are provided in Appendix B.

Figure 3 shows the results from the theoretical derivations of the number of slaughterhouses needed to be sampled, in a scenario based on the variability values illustrated above. In particular, the right panel shows the number of broilers to be sampled considering different numbers of sampled flocks, when 20 slaughterhouses are sampled. It should be noted that, with the chosen variability values, the sampling of only 10 slaughterhouses does not allow reaching the level of confidence and power set up for the specific configurations, hence the blank left panel in Figure 3. Moreover, Figure 3 shows that the number of broilers needed per flock decreases as the between flocks variability increases; this counterintuitive result is due to the fact that increasing flock variability also increases the effect size if the marginal prevalence is kept constant. Further details are available in Section 3.2 of the methodological report (EFSA, 2013).

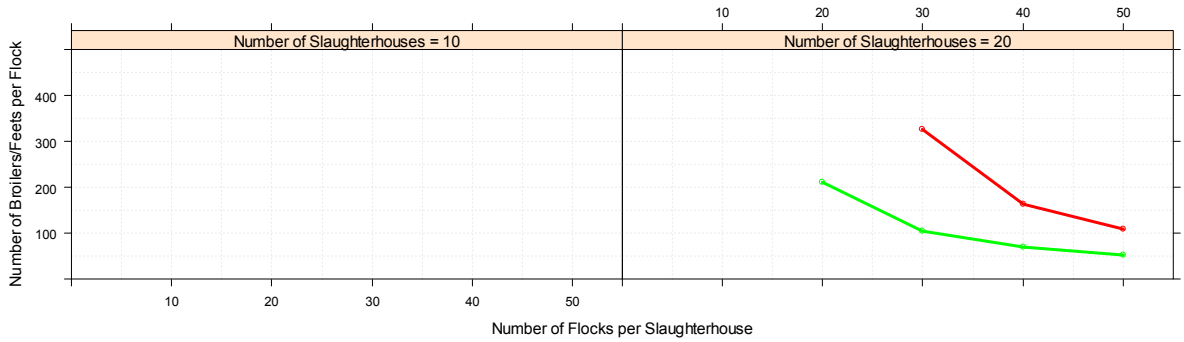


Figure 3: Number of broilers to be sampled per flock, considering different numbers of flocks sampled in 20 sampled slaughterhouses, variability at slaughterhouse level of 1 and two variability values for flocks (2.5 in full red line in red and 4.2 in full green line in green). The empty left panel indicates that, the sampling of only 10 slaughterhouses does not allow reaching the level of confidence and power set up for the specific configurations.

Finally, Figure 4 shows the total number of flocks and broilers/feet within flocks to be sampled according to the different powers to achieve and when slaughterhouses are treated as a fixed effect instead of as a random effect. This may happen when there are no prevalence differences between slaughterhouses or when all slaughterhouses in the MS are sampled.

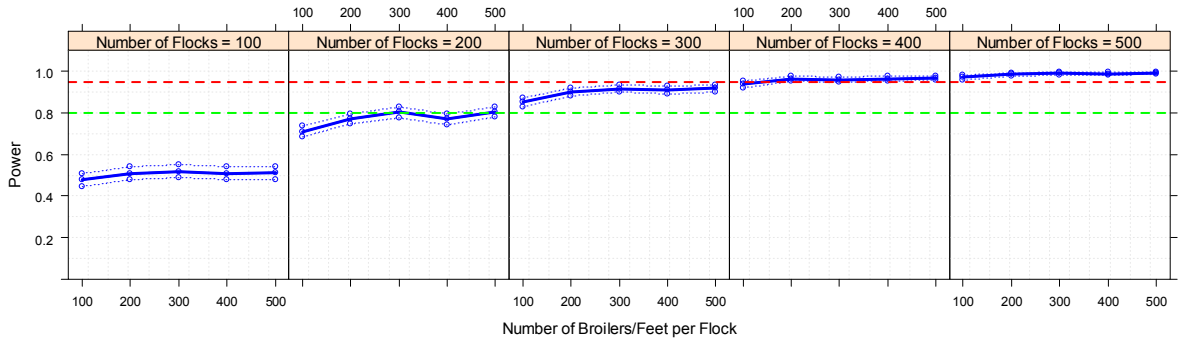


Figure 4: Power curve and total number of flocks and broilers/feet to be sampled for each simulated scenario, when all slaughterhouses are sampled, considering variability at flock level of 4.2 and effect size of 0.005 to detect population prevalence greater than 0.01. Green and red dash lines represent power levels of 0.8 and 0.95 respectively.

Simulations were run to calculate the power of the test in relation to variations for hock burn and foot pad dermatitis. In particular, for foot pad dermatitis, the assessment was done to estimate prevalence of the worst foot pad dermatitis condition- score 2 (the description of the scoring system is provided in Section 3.3.3 of this report); such sample size would be sufficient to detect also the foot pad dermatitis of scores 0 and 1. Furthermore, it is assumed that sample size estimation in relation to the other indicators of the draft data collection, outlined by the Commission's draft document, should in general be smaller given the nature of the indicator. More specifically, the other indicators to be collected contain information of the whole flock without introducing variability at the bird level. Hence, considering sample size formulas that account for this source of variability will be on the safe side.

3.2.2. Procedure to calculate sample sizes in MS with more than 20 broiler slaughterhouses

Given the following parameters considered appropriate for this data collection and on which the precision of the test depends:

- marginal prevalence of interest of 0.01 (1.0%)
- size of the accepted type I error α of 0.05 (5%)
- chosen statistical test power (1-type II error risk) of 80%
- effect size of 0.005 (0.5%), defined as the difference between the prevalence of 0.01 according to the null hypothesis and the true prevalence of 0.015 (1.5%)

the analysis of the above simulations indicates that, in this three level model it is necessary to sample at least 20 slaughterhouses in each MS to reach the necessary power of the test, assuming that the sample size represents less than 10% of the total number of slaughterhouses in the country (in the country there are at least 10 times the number of slaughterhouses sampled). If this assumption does not hold, a finite population correction applies which may slightly reduce the number of slaughterhouses to be sampled (see Appendix B).

In particular, Figure 1 shows that in order to detect a true prevalence of 1,5% with a variability at flock level of 2.5, it is necessary to sample at least 20 slaughterhouses, with 20 flocks sampled per slaughterhouse and 200 broilers/feet sampled per flock, when the number of slaughterhouses in the

country can be considered infinite (e.g. more than 10 times the sample size). Figure 2 shows that, for a variability at flock level of 4.2, 20 flocks per slaughterhouse and 300 broilers/feet per flock are needed. These results were compared to those of the above-mentioned theoretical formulas, used to estimate the necessary number of slaughterhouses to be sampled in a three level hierarchical population.

The results of the simulations and of the theoretical formulas were found to be consistent, after corrections due to the approximation used to derive the theoretical formulas. The correction of 10% increases the number of slaughterhouses when not using second order penalized quasi-likelihood estimations (PQL2) in the theoretical formulas, which are based on first order marginal quasi-likelihood estimations (MQL1). Full explanation is provided in Section 2.3.2 and 3.3.2 of the methodological report (EFSA, 2013).

The sample size values as computed from the theoretical formula (equation 1) described in Appendix B, including the above-mentioned increase of the number of slaughterhouses by 10%, as well as the sample size suggested above based on simulations, need to be adjusted for two further effects:

1. The total number of slaughterhouses in any MS is finite and always less than 200. If the number of slaughterhouses to be sampled is more than 10% of the total number of slaughterhouses in the country, a finite population correction factor (provided in appendix B) must be applied to the prevalence variation between slaughterhouses, which is part of the input for the sample size calculation. This correction must be done before the number of slaughterhouses to be sampled is computed and leads to a slightly smaller sample size.

2. Since flock size variation increases the sampling error of the prevalence estimate at MS level, the number of flocks sampled per slaughterhouse as assumed for computing the number of slaughterhouses to be sampled, must finally be increased by 20%. This correction must be done after the number of slaughterhouses has been determined. Technical details of this correction factor are reported in Appendix B.

3.2.3. Procedure to calculate sample sizes in MS with less than 20 broiler slaughterhouses

In the simulations where we can assume a two level model (flocks and broilers/feet, with all slaughterhouses sampled), Figure 4 shows that 300 is the total number of flocks needed to be sampled in the country with at least 100 broilers/feet tested for each flock, in order to achieve 80% power and a confidence of 95%. With a due increase of 20% of the number of flocks to account for possible flock size variations within slaughterhouses, a MS should sample a total number of 360 flocks with at least 100 broilers/feet tested for each flock.

As a consequence, in the case where there are less than 20 slaughterhouses in a MS, then all slaughterhouses should be sampled, estimating sample size for flocks and broilers/feet accordingly. This can be done using directly the values of Figure 4 and corrected by 20% for the flock size variations (360 flocks with 100 broilers/feet for each flock) when in the MS the sample size is smaller than 10% of the total number of flocks slaughtered (the total number of flocks is at least 10 times the number of flocks sampled).

In those MS where the sample size is higher than 10% of the total number of flocks slaughtered in the country (i.e. if the latter is less than 3600 flocks), then the sample size calculation has to be done using the theoretical formula because a finite population correction is again needed, in respect to the prevalence variation between flocks. In this case, the sample size calculation then proceeds as follows:

i) plug in, as number of slaughterhouses sampled, the total number of slaughterhouses in the country, and plug in “100” as number of birds sampled per flock;

ii) set the prevalence variability at slaughterhouse level at “0” (since there is no sampling error at slaughterhouse level anymore, as all slaughterhouses are sampled);

iii) multiply the prevalence variability at flock level with the finite population correction in Appendix B;

iv) compute the total number of flocks needed to sample in the MS;

v) increase this number of flocks by 30% to achieve agreement between formula and simulations.

In both cases, whether the number of flocks is based on the simulations or on the formula with corrections, two final steps are needed:

i) increase the total number of flocks by 20% to compensate the increased sampling error due to flock size variation within slaughterhouses (as already indicated above for the cases were simulation results can be directly applied);

ii) distribute the total number of flocks between all slaughterhouses in the MS, such that the sample size per slaughterhouse is proportional to their annual throughput. Further details are provided in Section 5 of this report and in the EFSA methodological report (EFSA, 2013).

The two above-mentioned procedures, with the mentioned values for variability and effect size as indicated above, were used to estimate the sample size needed for this data collection for each MS. In particular, Table 1 summarises the specific procedure that was applied for each MS.

Table 1: Procedure applied for the calculation of the sample size for each MS, depending on the number of broiler slaughterhouses in the country

Procedure applied for simple size calculation	**Number of slaughterhouse**			
	More than 20 slaughterhouses/MS		**Less than 20 slaughterhouses/MS**	
	Sample<10%	Sample>10%	All slaughterhouses sampled	
			360 flocks, 100 broilers/feet	
			Sample<10%	Sample>10%
Formula or simulation	Simulation	Formula	Simulation	Formula
Correction for finite population	Not needed	Yes $\left(\frac{n_3 - 10}{n_3 - 1}\right)$	Not needed	Yes $\left(\frac{n_2 - 200}{n_2 - 1}\right)$
Theoretical approximation correction	Not used	Yes (10% at the slaughterhouse level)	No (300 flocks given the simulation results)	Yes (at least 30% at flock level)
Correction for size variation	Not used	Yes (20% correction at slaughterhouse level)	Yes (20% correction at flock level, resulting in 360 flock)	Yes (20% correction at flock level)
MS	No MS	BE, CY, CZ, DE, ES, FR, PO, RO, UK	AT, DK, FI, NL, SE, SI	EE, IE , LT, LV

Based on the sample size calculation as explained above, the assessment of the 1% sample size for each MS is reported in detail in Section 4.2 and Appendix C of this report. For those MS where full data were available, EFSA has also provided possible sampling plans reported in Appendix C of this report.

3.2.4. Welfare indicators: background, definitions and use

Animal-based welfare indicators are the conditions to be examined in this specific data collection.

Traditionally, animal welfare legislation and also private standards have mainly been focused on input variables (hazard factors, as resources made available and management practices applied), by stating how animals (including poultry) should be housed, fed and managed. This is in general the basis for a clear and predictable legislation where the owner/stockperson can invest in a certain husbandry system, apply a certain stocking density or carry out certain management procedures knowing that this is all in compliance with the legislation. By emphasising the input variables, it has been clear that the aim of the legislation or standards has been to prevent animal welfare problems and unnecessary suffering. Nevertheless, it is generally accepted that although the legislation may be correctly applied this does not *per se* guarantee good animal welfare, as there may for example be aspects not foreseen or covered by the legislation. From a strict animal welfare perspective, it is the actual outcome (adverse effect) that matters to the animal, i.e. how the animal feels, its state of well-being and health. To cover also this aspect, animal welfare indicators – often animal based – have been developed and included in the legislation, in this case Council Directive 2007/43/EC.

An assessment of the animal-based measures used for the welfare of broilers is provided in the EFSA Scientific Opinion on the use of animal-based measures to assess the welfare of broilers (EFSA, 2012c). Moreover, the EFSA Statement on the use of animal-based measures to assess the welfare of animals provides general considerations on the use of animal-based measures (EFSA, 2012b).

Animal-based welfare indicators may be assessed on farm or at the slaughterhouse. For practical reasons, indicators that can be assessed or collected at the slaughterhouse are often preferred, as this facilitates the process and limits the need of a large number of on-farm visits by control staff or auditors. For a production system such as the modern broiler production, where flocks of birds are sent for slaughter up to 6 or 7 times a year, indicators assessed or collected at the slaughterhouse are particularly relevant, as this will lead to a continuous surveillance of bird welfare.

In practice, such indicators are already systematically collected in many countries. For example, the official meat inspection system in the MS often covers the recording of a relatively large number of parameters/diagnoses which can be used to asses not only food safety but also animal welfare, such as metabolic or infectious diseases leading to carcass rejection. Some MS (e.g. Sweden and Denmark) have been using slaughterhouse-based records of foot pad dermatitis as an indicator of broiler welfare for many years, and both foot pad dermatitis and hock burn are in many countries recorded by many of the food business operators, as a quality parameter. Both the competent authorities and the food business operators may also routinely be recording the proportion of birds being dead on arrival, to monitor animal welfare during transport but also for economic reasons. Furthermore, there is already a requirement in Council Directive 2007/43/EC for the producer to record the cumulative daily mortality of each flock of broilers during the entire rearing period, when the stocking density applied is higher than 33 kg/m^2. In addition to this, there may be private (industry- or retailer initiated) broiler welfare standards requiring similar types of animal-based measures to be recorded for any producers affiliated to a certain animal welfare scheme.

Although some indicators such as foot pad dermatitis and hock burn are being systematically used as mentioned above, it should, however, be noted that so far the scoring system varies among the different countries and studies, both in terms of categories of scoring (from 3 to 5 scores for both hock burn and foot pad dermatitis) and in terms of definition of each category. The Commission draft document outlining the data collection in EU slaughterhouses therefore also aims at harmonising the definition and use of foot pad dermatitis to allow for comparable results among countries.

In this context and although the assessment of the protocol itself and of the indicators used was not part of the technical assistance mandate to EFSA, EFSA highlighted the need to report where imprecise definitions, especially in the case of multinomial indicators such as foot pad dermatitis, can introduce errors (measurement errors) when collecting the outcome of interest during the data collection. Moreover, while for binomial indicators it is possible to identify prevalence, in case of multi-score indicators such as foot pad dermatitis with 3 scores categories, it is better to refer to

proportions for each indicator category. The following Section 3.3 addresses more in details the indicators used in the Commission draft document outlining the harmonised data collection.

3.3. Description of each indicator of the data collection

The 7 welfare indicators identified in the draft data collection are indicators either recorded by the owners and provided at the slaughterhouse (culling rate, cumulative daily mortality and standard mortality), or assessed directly at the slaughterhouse by operators or competent authorities (hock burn, foot pad dermatitis, total rejections and dead on arrival).

3.3.1. Indicators collected at farm level

3.3.1.1. Culling rate

The culling rate is collected by the owner and is transmitted to slaughterhouse with the flock. Culling is the act to intentionally kill chickens for two main reasons (EFSA, 2010): i) when birds are unsuitable to be farmed (e.g.: wrong sex, wrong body conformation, surplus to requirement) - this is the case of "voluntary culling"; ii) when, in animals suitable for production, animal welfare is seriously compromised (e.g. disease, sickness, injuries, lameness)- this is the case of "involuntary culling".

Depopulation (mass culling) has been used to control diseases such as avian influenza (Stegeman et al, 2004; Dubey et al, 2009). The method used for culling impacts animal welfare (van den Berg and Houdart, 2008; Alphin et al, 2010).

The culling rate is calculated by the division of the number of birds culled, either for disease or because of other reasons from the time the birds are placed in the house until the day the flock is sent to slaughter, by the number of chicken initially present in the house, multiplied by 100.

3.3.1.2. Cumulative daily mortality

The daily mortality is an indicator used to monitor animal health and welfare in farm or during experimental trials (Atencio et al, 2010; Balog et al, 1997; Xin et al, 1994).

The daily mortality rate, as proposed in the Commission draft document, is the number of chickens which have died in a house on the same day, including those that have been culled either for disease or because of other reasons, divided by the number of chickens present in the house on that day, multiplied by 100.

The cumulative daily mortality rate is the sum of all daily mortality rates.

It should be noted that the definitions of daily mortality rate and cumulative daily mortality, provided in the Commission draft document, are those provided in Council Directive 2007/43/EC. In particular, the cumulative mortality rate threshold can be used as a criterion to allow maximum stocking densities up to 42 kg/m^2.

3.3.1.3. Standard mortality

Mortality is an important indicator used to manage broiler health, welfare and performances. Standard mortality rate is the number of chickens that died in a house, including the culled ones, from the first day of rearing in that house until birds are sent to slaughter, divided by the number of chickens initially present in the house and multiplied by 100.

For the three above calculations (culling rate, cumulative daily mortality and standard mortality), the number of chicken initially present includes the chicken delivered for free by breeding companies.

Mortality itself does not directly reflect animal welfare but can impact welfare if we consider the way and the reason for animal dying, including culling. Culling of birds for welfare reasons, is a way of minimizing suffering, e.g. when broilers are suffering from leg disorders (Knowles et al, 2008).

Culling and mortality should be as low as possible; when animals are sick or injured, culling is the best way to prevent them from suffering. Even if the aim is to monitor and reduce overall mortality, the number of animals found dead should be considered separately from the number of animals culled, since culling of sick birds is desirable to reduce suffering. In this case, the ratio culled birds/birds found dead can be a useful indicator and can reflect a good management level (EFSA, 2010). On the contrary, a too high culling rate could depend on housing or management defaults that can lead to poor conditions of birds.

The main causes for mortality's variability include:

- Infectious or parasitic diseases such as infectious bursal disease (Igbokwe et al, 2012), colibacillosis (Won et al, 2009), necrotic enteritis (Dinev, 2010; Timbermont et al, 2009), botulism (Trampel et al, 2005), coccidiosis (Karaer et al, 2012) or association of several of them (Ganapathy et al, 2000);

- Metabolic disorders linked with fast growth rate principally (EFSA, 2010), such as ascites syndrome (Karlmar et al, 2013) and sudden death syndrome;

- Rearing management parameters such as: feed (Berhe and Gous, 2008), breed and sex (Casey et al, 1989; EFSA, 2010).

3.3.2. Indicators collected routinely at the slaughterhouse

3.3.2.1. Dead on arrival

Dead on arrival is recorded at slaughter, before shackling. The dead on arrival rate is the number of chickens in the flock arriving dead at the slaughterhouse, after transportation, divided by the number of chickens transported, multiplied by 100.

Annex III of Directive 2007/43/EC, point 1.2, provides for the number of broilers dead on arrival to be recorded under the supervision of the official veterinarian.

Dead on arrival can depend on several factors, including for example transport duration and climatic condition (Silva et al, 2011), catching system, and stocking density in crates (Chauvin et al, 2011; Ekstrand, 1998). Moreover, it should be noted that in slaughter-plants where gas stunning is used, the level of dead on arrival recorded could be underestimated as broilers are transferred directly in the crates from the truck to gas stunning system.

3.3.2.2. Total rejection rate

Total rejection rate is the number of carcasses rejected per flock indicating poor welfare conditions, such as abnormal levels of contact dermatitis, parasitism or systemic illness in the house. This definition excludes carcasses rejected due to slaughter process.

Total rejections are recorded on the slaughter line at the post mortem inspection.

Total rejection includes: ante mortem rejects (culls/runts), abnormal colour/fevered, emaciation, ascites/oedema, bruising/fractures, joint lesions/arthritis, cellulitis, dermatitis, tumours/nodules, respiratory disease (Airsacuulitis), salpingitis, hepatitis, perihepatitis and other (eg jaundice). In most slaughterhouses where there is no cutting plant, if any part of the carcass is condemned, the whole carcass is usually condemned. However, in slaughterhouses including a cutting plant, the total rejections could be lower than in case of slaughterhouses without a cutting plant.

It should be noted that the list of total rejections' causes of the Commission draft document includes also dead on arrival. As dead on arrival is to be recorded anyway as an individual indicator, it could be deleted from this list.

Rejection rate can vary with several parameters, including the catching system (Ekstrand, 1998). Causes for rejection are often infectious diseases (Yogaratnam, 1995), cellulite, contusion/ fractures and bruises (Santana et al, 2008). Lupo et al (2010) studied, on 404 broiler chicken flocks, risk factors for both condemnation categories (infectious and traumatic), and the relative impact of the different production stages (farm structure and routine husbandry practices; on-farm flock history and characteristics; catching, transport and lairage conditions; slaughterhouse and inspection features) on the whole condemnation rate. Results showed that significant factors were either specific to one condemnation category or related to both categories, and each of the explanatory blocks was involved in the explanation of infectious and traumatic condemnation rates. On-farm flock information explained 40% of the overall condemnation process whereas the other explanatory blocks had similar relative impacts. On the same sample the condemnation rate was 87 per 10,000 birds slaughtered (95% confidence interval 79 to 95 per 10,000) but differed significantly according to the type of poultry produced (standard, light, heavy or certified). The main reasons for condemnation were emaciation and congestion, with rates of 30 and 22 per 10,000 birds slaughtered, respectively. Congestion was significantly associated with arthritis and ascites, whereas infected skin lesions were associated with bruises and abnormalities of colour, odour or conformation (Lupo et al, 2008).

3.3.3. Indicators collected at slaughterhouse by specific scoring: foot pad dermatitis and hock burn

The Commission draft document, outlining a harmonised data collection, defined foot pad dermatitis and hock burn as follows:

Foot pad dermatitis:

Place of collection of data: after scalding on slaughter line or from feet collected in a box from the line.

Scores: 0,1 and 2[7]

Score 0: healthy foot
Area affected: none or very small
Lesion description:
- very small superficial lesions (1-2 mm)
- slight discolouration in a limited area
- mild hyperkeratosis
- completely healed scar

Score 1: foot with mild lesion
Area affected: does not extend over entire plantar pad
Lesion description:
- substantial discoloration of the foot pad
- superficial lesion, no ulceration
- dark papillae with no ulceration

Score 2: foot with severe lesion
Area affected: greater surface of plantar pad usually affected, sometimes with lesions on toes.
Lesion description:
- deeper lesions with ulceration, sometimes haemorrhage
- scabs of significant size

[7] Assessment rules similar as those used by DEFRA.

- severely swollen pad

Hock burns:

Place of collection of data: Slaughter line after scalding or after plucking

Scores: Scores 0, 1 (absence or presence)

Score 0: no lesions or very small
Area affected: none or very small
Lesion description: absence or very small (up to 2 mm) superficial lesions, slight discoloration in a limited area (up to 2 mm)

Score 1: presence of lesions
Area affected:
Lesion description: presence of superficial lesions (more than 2 mm), discoloration in an area bigger than 2 mm, some brown to black discoloration of the hock.

According to the 2010 EFSA Scientific opinion (EFSA, 2010), contact dermatitis (skin lesions of the breast, hock and feet) is a widespread problem in European broiler production even if the incidence is highly variable (Bessei, 2006; Berg 2004). In severe cases the erosions develop into ulceration with inflammatory reactions of the subcutaneous tissue. The lesions can become infected with a variety of bacteria and can even lead to joint inflammation. Such lesions can cause pain, whether infected or not, which constitutes a welfare issue. Management practices (litter, air quality, feed) as well as genotype, housing, age of birds (EFSA, 2010) and seasonal effect (de Jong et al, 2012) have an impact on their prevalence. The visual inspection at slaughterhouse is an ideal point where animal-based measures such as contact dermatitis can be visually detected and recorded. Allain et al (2009), studied broilers lesions at the slaughterhouse on 55 flocks. The presence of contact dermatitis was related to litter quality and several management factors and the type of lesions was related to the genotype. On the basis of correlations between lesions, the authors recommend an assessment of broiler chicken welfare based on: foot pad dermatitis, hock burn, breast burn, scratches and breast blister observations.

As explained in the 2012 EFSA Scientific Opinion (EFSA, 2012c), a system of routine recording of foot pad dermatitis in broilers was developed in the mid-1990s in Sweden and later adopted in Denmark, and is now applied at all broiler slaughterhouses in these countries. The system (Berg and Algers, 2004) is based on the visual examination of a systematic random sample of 100 single broiler feet per flock with a score from 0 (no lesions) to 1 (mild, superficial lesions, discolouration) or 2 (severe lesions, ulcers). These scores are then weighted and summarised to give a total flock score, which is used as a direct indicator of foot pad dermatitis and an indirect indicator of litter quality and bird management. A similar scoring system using a five-point scale with clear visual description of the lesions associated and histologically validated has recently been proposed (Michel et al, 2012). This system could easily be translated in a 3 points-scale with scores from 0 (no lesion or enlarged scales), to 1 (hypertrophic and hyperkeratotic scales covered by yellowish to brownish exudates), to 2 (depressed lesion, ulceration, with or without dark thick adherent crust). These two systems have been taken into account by EFSA to refine the definitions of foot pad dermatitis in the Commission draft document proposing a harmonised data collection.

3.3.4. Revised definitions of foot pad dermatitis and hock burn as proposed by EFSA

For each indicator and in particular for indicators with score categories, the definition is very important in the context of an EU survey. Indeed, when aiming at comparing prevalence of these parameters among MS, the overall exercise must be subjected to as less bias as possible. The very precise and clear definition of each criterion is needed in order to prevent for any bias of classification of a bird or a flock. Comparison among MS will not be possible if data collection is not strictly done with the same criteria definition, especially when the criteria need calculation or classification

regarding macroscopic aspect. Moreover, it is important that indicators are assessed in a consistent way by skilled staff; for this purpose instructions are developed and distributed and people are provided the appropriate training.

As the definitions of hock burn and foot pad dermatitis presented in the Commission draft document could leave some margin of interpretation and therefore differences in classification, EFSA has proposed more precise definitions of the different scores, based on the severity of the lesions.

The revised definitions are proposed below:

Foot pad dermatitis:

Score 0: healthy foot
Lesion description:
- No lesion or very small superficial lesions(diameter up to 1-2 mm)
- slight discoloration
- or/and mild hyperkeratosis
- or/and completely healed scar

Score 1: foot with mild lesion
Lesion description:
- substantial (diameter larger than 2 mm) discoloration of the foot pad
- or/and mild lesion, no (or very small: 1-2 mm) ulceration or scab
- or/and dark (brown) papillae with no ulceration

Score 2: foot with severe lesion
Lesion description:
- severe deeper lesions with ulceration, with or without hemorrhage
- scabs
- severely swollen pad

Hock burn:

Score 0: no lesions or very small superficial lesion
Lesion description:
- No lesion, slight discoloration or very small superficial lesion (diameter up to 2 mm)

Score 1: presence of lesions
Lesion description:
- superficial lesions (diameter more than 2 mm) or more severe lesions of any size
- brown to black discoloration of the hock
- and/or deeper lesions

It should also be noted that the exact moment in which the indicator is assessed (for foot pad dermatitis, after scalding or from feet collected in a box from the line) can influence the sampling protocol used; a general approach is provided in Section 5 of this report.

3.3.5. Practical considerations on the indicators included in the draft data collection

When considering indicators to be collected for a monitoring or surveillance scheme, it is important to consider not only their reliability but also their practicability.

When it comes to mortality indicators for example, the Commission draft document foresees the recording of both cumulative daily mortality and standard mortality. It should however be noted that in statistical terms, the two parameters are related and that the one can be calculated from the other.

The statistical interpretation of the two indicators and their correlation are provided in Appendix H of this report.

4. RESULTS OF THE ASSESSMENT

4.1. Data used for the analysis

In the methodological steps of the technical assistance, case studies and pilot projects carried out in UK and France on the collection of foot pad dermatitis and hock burn (Haslam et al, 2008; Allain et al, 2009 and a pilot study recently conducted; EFSA, 2013) were used to fit the model and to estimate the specific between-slaughterhouse and between-flock variability, which were found to be significantly different from 0. The description of how scenarios for theoretical computation and simulation were set up, based on those case studies, can be found in Section 2.2 of the EFSA methodological report (EFSA, 2013).

The assessment of the sample size by MS required the knowledge of the total number of broiler flocks slaughtered in the MS in a year time. Moreover, where information was fully available also in relation to slaughterhouses and number of broilers slaughtered, it was possible to suggest possible sampling schemes by MS, set on the approach described in Section 5 of this report.

The analysis was carried out using data obtained directly by the MS for the purpose of this specific technical assistance mandate and upon request.

Chief Veterinary Officers were requested to nominate an official reporting officer for their country and to provide EFSA with two levels of data:

i) data at country level, including the total number of broiler slaughterhouses, the total number of broiler flocks and the total number of broilers slaughtered in a given year. These were the minimum data requested to be able to carry out the assessment;

ii) data at individual slaughterhouse level, including the regional location of the slaughterhouse, the throughput in terms of broiler flocks and broilers slaughtered per year, referring to the same year indicated in the data at country level. Where fully available, this information was used to design sampling plans, taking into account the slaughterhouses annual throughput.

23 MS reacted to the data sharing agreement, in particular:

- 19 MS provided the necessary data at country level;
- 15 provided full data also at slaughterhouse level;
- for 2 MS information expressed in number of broiler flocks was not available;
- 1 MS has no broiler slaughterhouses.

It should be noted that some countries did not record the number of flocks slaughtered but rather the number of batches or birds slaughtered. In these cases and with the sampling unit being the flock, it was not possible to convert the number of batches or birds in flocks, hence the assessment of the sample size was not carried out.

4.2. Assessment of the sample size by Member State

The assessment of the sample size by each MS, for which the necessary data were available, was based on the procedures illustrated in Sections 3.2.2 and 3.2.3 of this report. The assessment was performed only for 19 MS for which the data submitted allowed the analysis.

Depending on the production of the MS, first the necessary sample size or possible combinations of sample sizes (number of slaughterhouses, number of flocks per slaughterhouse and number of broilers/feet per flock) were calculated. Secondly, in order to assess if the chosen 1% sample of flocks would ensure to achieve the objectives of the data collection, the number of flocks needed (computed from the sample size calculation) was compared to the total number of flocks slaughtered in a year in each MS.

In MS having more than 20 broiler slaughterhouses, the necessary sample size was calculated using the three level formula reported in Appendix B and considering different possible combinations of number of slaughterhouses to sample, number of flocks within each slaughterhouse and number of broilers/feet within each flock. In particular, firstly the possible representative combinations of number of flocks per slaughterhouse/broilers and feet per flock were selected and then the number of slaughterhouses to be sampled computed accordingly. In a second stage, the most suitable combination that better fitted with the proposal of the Commission draft document (total sample of minimum 200 flocks and maximum 500 flocks to be tested per MS and with 300 broilers per flock), was identified.

Appendix C presents the individual tables by MS showing the full sufficient sample sizes for each given combination and, in correspondence of each combination, the assessment of the 1% flocks. The combination that EFSA considered the most suitable (a total of 24 flocks to be tested per slaughterhouse and 300 broilers/feet to be tested per flock) is always reported as combination 1 in all tables, where the number of slaughterhouses to be sampled changes from one MS to another. The full tables are provided in this report as they could serve as a practical tool to risk managers to set the sampling protocols according to the resources available and the number of slaughterhouses that can be sampled.

This procedure was applied to Belgium (BE), Cyprus (CY), Czech Republic (CZ), France (FR), Germany (DE), Poland (PO), Romania (RO), Spain (ES) and UK. Moreover, for those MS where full data were available also at slaughterhouse level and for the above-recommended combination number of flocks/broilers and feet, EFSA has also provided a sampling plan to randomly distribute the necessary sample of slaughterhouses, based on the individual slaughterhouse throughputs. The sample weighting was calculated on the total number of broilers slaughtered by each slaughterhouse. These sampling plans are provided in Appendix C together with the tables for each MS.

The analysis of the above-mentioned assessment shows that if only 100 broilers/feet are tested per flock, the necessary total number of flocks to be tested in each MS is always higher than 500. Moreover, given the variability parameters and conditions considered in this assessment, it also shows that hock burn and foot pad dermatitis could be assessed on the same sample broilers/feet and that no differentiation is needed.

In this context, for 5 MS (CY, CZ, DE, PL and RO) 1% of the total number of flocks slaughtered in one year time is not a sufficient sample to ensure the objectives of the data collection, according to the parameters chosen to calculate the sample size needed. For 3 MS (BE, ES and UK), the representativeness of the 1% sample depends on the combination number of flocks/broilers and feet chosen. For these 3 MS when the recommended combination (24 flocks per slaughterhouse and 300 broilers/feet for each flock) is used, 1% of the total number of flocks slaughtered in one year time is a sufficient sample to ensure the objectives of the data collection according to the parameters chosen. For 1 MS (FR), when any combination is used, 1% of the total number of flocks slaughtered in one year time is a sufficient sample to ensure the objectives of the data collection, according to the parameters chosen.

In MS having less than 20 broiler slaughterhouses and where all of them need to be sampled in a two level model (Austria (AT), Denmark (DK), Estonia (EE), Finland (FI), Ireland (IE), Latvia (LV), Lithuania (LT), The Netherlands (NL), Slovenia (SI) and Sweden (SE)), two different procedures were followed depending on whether:

i) the total number of flocks sampled in the MS (360 flocks) was lower than 10% of the total number of flocks slaughtered (at least 3600 flocks slaughtered in the MS- this procedure was also applied to countries with slightly less than 3600 flocks slaughtered per year);

ii) the total number of flocks sampled was higher than 10% of the total number of flocks slaughtered (less than 3600 flocks slaughtered in the MS).

The first procedure, applied to AT, DK, FI, NL, SE and SI, allowed the calculation of the total number of flocks to be sampled in the MS based on the results of the simulation in Figure 4, corrected by 20% for the flock size variations (360 flocks with 100 broilers/feet for each flock).

The second approach, applied to EE, IE, LT and LV, allowed the calculation of the representative total number of flocks to be sampled in the MS based on the three level formula, with the total number of slaughterhouses in the country to be sampled and with 100 broilers/feet to sample for each flock, and as described in Section 3.2.3 of this report.

In both cases for these 10 MS (AT, DK, EE, FI, IE, LT, LV, NL, SE and SI), 1% of the total number of flocks slaughtered in one year time is not a sufficient sample to ensure the objectives of the data collection, according to the parameters chosen. The results by each MS are provided in Tables 2-33 in Appendix C.

In addressing this technical assistance mandate, the sample size was calculated in order to detect a marginal prevalence of the conditions of 0,01, with an effect size of 0,005, considering 80% power of the test and 95% level of confidence. In order to be able to cover the range of variability level in all MS and among the parameters illustrated in the previous sections, it was considered appropriate to use the higher variability at flock level for hock burn and foot pad dermatitis of 4.2, when using the two level model. In case of the three level model, the variability considered at slaughterhouse level was 1 and the variability used at flock level was 2.5, as shown in the results obtained from the case studies reported in the methodological report and due to the inclusion of an extra level of variation (slaughterhouse level).

It should be noted, however, that the three level formula provided in Appendix B allows calculating directly the sample sizes, depending on the variability parameters that are plugged in the formula.

Moreover, the sample size strictly depends on the level of prevalence to detect and on the effect size chosen. Table 34 in Appendix D shows the sample sizes needed for different combinations of number of slaughterhouses/flocks/broilers and feet, to detect different levels of prevalence under the null hypothesis (1%, 5% and 10%), when considering two levels of relative effect sizes of 1.5 and 2 (relative to the prevalence under the null hypothesis of 1%, 5% and 10%) and given 2 different levels of flock variability (2.5 and 4.2).

Finally, for 14 MS where complete data were available (AT, BE, CY, CZ, DK, ES, FI, FR, IE, LV, LT, RO, SI and UK), possible sampling plans, built on the approach explained in the following Section 5 of this report, are proposed and reported in Tables 3-34 in Appendix C. In particular, for those MS where a three level model was used for the assessment, EFSA has provided the list of slaughterhouses to be sampled, based on the weighted random selection of slaughterhouses depending on their throughputs.

For MS with less than 20 slaughterhouses where all slaughterhouses are to be sampled, EFSA has provided the allocation of numbers of flocks to be tested in each slaughterhouse, depending on their annual throughput.

5. DESIGNING A SAMPLING PLAN FOR THE DATA COLLECTION OF WELFARE INDICATORS AT BROILER SLAUGHTERHOUSES WITHIN EACH MS

Based on the methodological approach developed for this mandate, this section proposes a general approach and steps to follow in order to design a sampling plan in each MS. As the objective of the data collection is to assess prevalence of given welfare conditions at the MS level, the flocks to be sampled should be farmed in the given MS, according to Council Directive 2007/43/EC.

This approach could be followed, once the needed sample size has been calculated.

5.1. Information required to produce a sampling plan

For each MS, a list of all the broiler slaughterhouses, with slaughterhouse specific annual kill (total number of broilers slaughtered annually) is required.

5.2. Proposed sampling approach to select slaughterhouses and number of flocks per slaughterhouse

5.2.1. Slaughterhouse Selection

For MS with more than 20 slaughterhouses, each slaughterhouse within a MS is assigned a probability of selection that is proportional to the total number of broilers it slaughters per year. A weighted random selection of slaughterhouses is then taken for each MS.

Where MS have provided EFSA with all the data also at slaughterhouse level, EFSA has provided the list of slaughterhouses to sample, based on this approach. The plans are included in Appendix C, together with all the results for each MS.

The theoretical maximum number of slaughterhouses to be selected per MS is 20, except where the total number of slaughterhouses in a MS is less than 20, in which case all slaughterhouses within a MS are to be sampled.

Further, for the MS with more than 20 slaughterhouses but less than 200 in total (where the sample size is more than 10% of the total population - this being the case for the MS so far analysed by EFSA) , the sample may be reduced by applying a 'correction factor' to that part of the sampling error which is due to prevalence variation between slaughterhouse.

More specifically, the outcome variance at the slaughterhouse level must be multiplied with the following finite population correction factor:

$$\frac{n_3-10}{n_3-1}$$

where n_3 is the total number of slaughterhouses in a MS, and the number of slaughterhouses to be sampled should then be calculated based upon this correction factor in conjunction with the following equation, given the number of flocks to sample and the number of broilers/feet to sample for each flock:

$$n_3 = \left(Z_{1-\beta} + Z_{1-\alpha}\right)^2 \times \left(\frac{\sigma_S^2 + \frac{\sigma_F^2}{n_2} + \frac{\sigma_e^2}{n_2 n_1}}{\left(\beta_{0A}^C - \beta_{00}^C\right)^2} \right)$$

Where $\sigma_e^2 = [\pi^C(1-\pi^C)]^{-1}$ is the variance at the broiler level, given $a = b = 0$ and using the true β_{0A}^C to compute π^C, and $\Delta = \beta_{0A}^C - \beta_{00}^C$ is the effect size.

The full explanation of the above equation is provided in Appendix B of this report and in Section 2.3.2 of the EFSA methodological report (EFSA, 2013).

The value "10" is used within the correction factor equation instead of 20, purely to maintain a conservative approach to the sample size as 10 is the minimum number of slaughterhouses to be sampled per MS and thus gives the weakest downward correction of the sample size needed.

5.2.2. Flock Selection

The total number of flocks to be sampled per MS will be a fixed number and this total will be MS specific, also based upon the above correction factor, which affects the number of slaughterhouses to be sampled.

The total number of flocks to sample for each MS has been calculated based upon the data provided by MS and is given in the tables included in Appendix C of this report, except where the data requested by EFSA were not available.

Within MS, the number of flocks to be sampled should be distributed evenly across all the MS's slaughterhouses, i.e. each slaughterhouse should have the same number of flocks sampled.

However, if the MS has less than 20 slaughterhouses in total, then all slaughterhouses in the MS must be sampled and the total number of flocks to sample within the MS as a whole will be 360; this number should, differently to the above, be allocated to slaughterhouses in proportion to the total number of flocks processed per year. The estimate of 360 flocks per MS for those with less than 20 slaughterhouses is based upon the results of two-level simulations reported in Sections 3.2.3 and 4.2 of this report. In the case that the number of flocks to be sampled is more than 10% of the total number of flocks slaughtered in the MS, then the sample may be reduced by applying the 'correction factor' to that part of the sampling error which is due to prevalence variation between flocks:

$$\frac{n_2 - 200}{n_2 - 1}$$

where n_2 is the total number of flocks in a MS, following the same principle as for slaughterhouse level.

The allocation of numbers of flocks to slaughterhouse for those MSs with less than 20 slaughterhouses is given in the tables by each MS provided in Appendix C, except where the data requested by EFSA were not available.

5.2.3. Broilers (and feet) Selection

For each MS with more than 20 slaughterhouses, 300 birds should be sampled from each flock selected for sampling (except in the case where the MS elects to choose another scheme from the tables in Appendix C).

For each MS with less than 20 slaughterhouses, 100 birds should be sampled from each flock selected for sampling.

These numbers are again based upon the results of simulations and equations reported in Sections 3.2.2 and 3.2.3 and 4.2 of this report.

5.2.4. Allocation of Flock Sampling within Slaughterhouse

In MS with more than 20 slaughterhouses and where 24 flocks are to be sampled according to the recommended sampling combination, sampling of 24 flocks within a slaughterhouse should be spread evenly across the months of the year (2/month); the day of the month and the flock within the day being selected at random.

Within the flock the number of birds, as outlined above, should be sampled at random from the processing line. Alternatively, the flock could be divided into three: first third, second third and last third, calculated from the line speed and known size of the flock, where one third of samples are collected at a random time within each of the thirds of the flock.

5.2.5. Hock and Foot Sampling within Flocks at Slaughterhouse

It has been found in the past that the best place, and the simplest place, to collect hock burn and foot pad dermatitis data is at the **first post-mortem inspection point** on the slaughter-line.

The first post-mortem inspection point is usually at the point where the carcasses are transferred from the poultry slaughter shackle line to the separate evisceration shackle line. On the slaughter shackle line, the birds are suspended from their feet. At transfer, the feet are cut off, and on the evisceration line carcasses are suspended by the hocks. At this point the carcasses are very clean and any blemishes are much more visible. Some slaughterhouses actually have video cameras at this point that can automatically monitor the feet for foot pad dermatitis.

At the point where the carcasses are transferred from one shackle line to the other, the hocks are generally facing and easily seen, and the feet can either be scored on the line before they are cut off or a sample of feet picked off the conveyor one at a time after they have been cut off. The feet are transferred away at such a rate that if only one foot is chosen at a time, there is no chance of picking up a second foot from the same bird. If feet are collected in a bin, care should be taken to only sample either left or right feet so as not to double sample from birds.

5.2.6. Collection of the other indicators

Variables such as dead on arrival, food chain information (FCI), culling rate, cumulative daily mortality and standard mortality are usually collected from the slaughterhouse office at the end of the visit. All these data are usually kept and collated there and easily accessible by working with the slaughterhouse system.

6. THE FUTURE DATA COLLECTION: POSSIBLE FUTURE SUBMISSION OF DATA TO EFSA THROUGH THE EFSA DATA COLLECTION FRAMEWORK- THE DATA MODEL

If an EU harmonised data collection will be implemented in the future in the MS, the possibility exists that MS will be requested to submit the data collected directly to EFSA for their collection and analysis at EU level. In view of this possibility, EFSA proposes in Appendix F a possible data model to submit the information to EFSA in a standardised and controlled way through the EFSA Data Collection Framework (DCF).

The data model was prepared taking into account the information to be collected according to the Commission draft document as well as additional information needed to allow analysis of the data and generation of preliminary epidemiological hypothesis on the correlation of risk factors and recorded welfare conditions. The data model is presented in Appendix G of this report.

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

Based on the methodology developed to reply to this technical assistance mandate and on the information provided by MS, the assessment of the 1% sample size was performed only for 19 MS for

which the data submitted allowed the assessment. The assessment was performed based on the parameters chosen and illustrated in this report (prevalence to be detected of 1%, effect size of 0.5%, power of the test of 80% with 95% level of confidence, variability at slaughter level of 1 and at flock level of 2.5 for the three level model, and variability at flock level of 4.2 for the two level model). Although the chosen variability parameters were considered appropriate in order to follow a precautionary approach, it should be noted that other values could be used for the calculation. It was found that for 15 MS (AT, CY, CZ, DE, DK, EE, FI, IE, LT, LV, NL, PL, RO, SE and SI), 1% of the total number of flocks slaughtered in one year time is not a sufficient sample to assess 1% prevalence of the given conditions. For 3 MS (BE, ES and UK), the 1% sample is sufficient depending on the combination number of flocks/broilers and feet chosen. For 1 MS (FR), when any combination is used, 1% of the total number of flocks slaughtered in one-year time is a sufficient sample to assess a prevalence of 1%.

For MS with more than 20 broiler slaughterhouses and where a three level model could be used, several scenarios of combinations (number of slaughterhouses/number of flocks per slaughterhouse/number of broilers and feet per flock) that can ensure the detection and assessment of 1% prevalence in the population are given in the tables provided in Appendix C.

Among all the possible above-mentioned combinations, the one that was considered the most appropriate by EFSA is to sample 24 flocks per slaughterhouse with 300 broilers/feet per flock. Such combination seems to be feasible, allowing an easy monthly distribution of flocks in slaughterhouses over the year, in order to consider also seasonality effects. Moreover, such combination would allow sampling a maximum of, or slightly above, 500 flocks per MS and a maximum of 300 birds per flock, as foreseen by the Commission draft document for a harmonised data collection. The corresponding number of slaughterhouses to be sampled in each MS for this combination is provided in the corresponding Tables of Appendix C.

For the MS with less than 20 slaughterhouses, all slaughterhouses should be sampled. The total number of flocks to be sampled in the country is 360 flocks, with 100 broilers/feet sampled per each flock, if the total number of flocks in the country is at least 10 times the number of flocks sampled.

If the total number of flocks in the country is less than 10 times the number of flocks that should be sampled, then a finite population correction should be applied, decreasing the number of flocks. Full results of this assessment are provided in Appendix C.

To ensure representativeness of the sample, the number of slaughterhouses needed in each MS should be selected randomly. The number of flocks/slaughterhouses should be evenly sampled during the year (in the recommended scheme: 2 flocks per slaughterhouse/month) by randomly selecting 2 days of sampling among working days in the month and each day, by randomly selecting the rank of slaughtering of the flock within the day. At each month, the randomization should be run again.

Within a flock, 3 subsamples (1/3 of the total sample) where birds should be randomly selected, should be taken in the first, second and last third of the flock to assess hock burn. When feet are cut, similarly, 3 subsamples (of 1/3 of the total sample) where feet should be randomly selected should be taken in the first, second and last third of the flock.

RECOMMENDATIONS

The sample size, identified in the Commission draft document proposing a harmonised data collection, should be increased as it seems insufficient to detect prevalence of 1% (with effect size of 0.5%, power of the test of 80% with 95% level of confidence, variability at slaughter level of 1 and at flock level of 2.5 for the three level model, and variability at flock level of 4.2 for the two level model).

The sample sizes by MS sufficient to achieve the objectives of this data collection could be calculated using the three level theoretical formula provided in Annex B of this report and following the steps there described.

For the MS that provided information to EFSA, the sample size could be based on the assessment and possible sample combinations provided in Appendix C.

In order to allow analysis of the results, all data collected should be registered in a standardized way and in a proper database. If a harmonised data collection will be implemented in the future and in case the data will be collected by EFSA, data should be submitted through the EFSA Data Collection Framework (DCF), following the data model provided in Appendix G.

The systematic recording of outcome measures can help in determining trends over time. It is crucial that the chosen indicators are valid and that can be measured reliably, reducing as much as possible measurement errors. In this context, the definitions and scoring system of foot pad dermatitis and hock burn should be as precise as possible in order to avoid different interpretations.

REFERENCES

Allain V, Mirabito L, Arnould C, Colas M, Le Bouquin S, Lupo C and Michel V, 2009. Skin lesions in broiler chickens measured at the slaughterhouse: Relationships between lesions and between their prevalence and rearing factors. British Poultry Science, 50(4), 407-417.

Alphin RL, Rankin MK, Johnson KJ, Benson ER, 2010. Comparison of water-based foam and inert-gas mass emergency depopulation methods. Avian Diseases, 54 (Suppl. 1), 757-762.

Atencio JL, Fernández JA, Gernat AG, Murillo JG, 2010. Effect of pine wood shavings, rice hulls and river bed sand on broiler productivity when used as a litter. International Journal of Poultry Science, 9(3), 240-243.

Balog M, Bayyari GR, Rath NC, Huff WE and Anthony NB, 1997. Effect of intermittent activity on broiler production parameters. Poultry Science, 76(1), 6-12.

Berg C, 2004. Pododermatitis and hock burn in broiler chickens. In: Measuring and auditing broiler welfare. Eds Weeks CA and Butterworth A. CABI, Wallingford, UK, 37-49.

Berg C and Algers B, 2004. Using welfare outcomes to control intensification: the Swedish model. In: Measuring and auditing broiler welfare. Eds Weeks CA and Butterworth A. CABI, Wallingford, UK, 223-229.

Berhe ET, Gous RM, 2008. Effect of dietary protein content on growth, uniformity and mortality of two commercial broiler strains. South African Journal of Animal Science, 38, 4, 293-302.

Bessei W, 2006. Welfare of broilers: a review. Worlds Poultry Science Journal, 62, 455-466.

Casey NH, Smith GA and Crosley RI, 1989. The influence of breed and sex on the incidence of mortalities and skin tears in broiler carcasses. Journal of the South African Veterinary Association, 60, 2, 102-103.

Chauvin C, Hillion S, Balaine L, Michel V, Peraste J, Petetin I, Lupo C and Le Bouquin S, 2011. Factors associated with mortality of broilers during transport to slaughterhouse. Animal, 5(2), 287-293.

de Jong I C, van Harn J, Gunnink H, Hindle V A and Lourens A. 2012. Footpad dermatitis in Dutch broiler flocks: Prevalence and factors of influence. Poultry Science, 91(7), 1569-1574.

Dinev I, 2010. Enzootic outbreak of necrotic gastritis associated with Clostridium perfringens in broiler chickens. 2010. Avian pathology, 39(1), 7-10.

Dubey SC, Nagarajan S, Tosh C, Bhatia S, Krishna LAL, 2009. Avian influenza: a long-know disease and its current threat. Indian Journal of Animal Sciences, 79(2), 113-140.

EFSA Panel on Animal Health and Welfare (AHAW), 2010. Scientific Opinion on the influence of genetic parameters on the welfare and the resistance to stress of commercial broilers. EFSA Journal 2010;8(7):1666, 82 pp. doi:10.2903/j.efsa.2010.1666

EFSA Panels on Biological Hazards (BIOHAZ), on Contaminants in the Food Chain (CONTAM), and on Animal Health and Welfare (AHAW), 2012a. Scientific Opinion on the public health hazards to be covered by inspection of meat (poultry). EFSA Journal 2012;10(6):2741, 179 pp. doi:10.2903/j.efsa.2012.2741

EFSA Panel on Animal Health and Welfare (AHAW), 2012b. Statement on the use of animal-based measures to assess the welfare of animals. EFSA Journal 2012;10(6):2767, 29 pp. doi:10.2903/j.efsa.2012.2767

EFSA Panel on Animal Health and Welfare (AHAW), 2012c. Scientific Opinion on the use of animal-based measures to assess welfare of broilers. EFSA Journal 2012;10(7):2774, 74 pp. doi:10.2903/j.efsa.2012.2774

European Food Safety Authority, 2012d. Scientific and technical assistance on *Echinococcus multilocularis* infection in animals. EFSA Journal 2012;10(11):2973, 22 pp. doi:10.2903/j.efsa.2012.2973

European Food Safety Authority, 2013. Sample Size Considerations for Hierarchical Population. EFSA Journal 2013;11(6):3292, 68 pp. doi:10.2903/j.efsa.2013.3292

Ekstrand C, 1998. An observational cohort study of the effects of catching method on carcase rejection rates in broilers. Animal Welfare, 7(1), 87-96.

Ganapathy, K., Salamat, M.H., Lee, C.C., Johara M.Y., 2000. Concurrent occurrence of salmonellosis, colibaccillosis, and histomoniasis in broiler flock fed with antibioticfree commercial feed. Avian Pathology, 29(6), 639-642.

Haslam, S. M., Knowles, T. G., Brown, S. N., Wilkins, L. J., Kestin, S. C., Warriss, P. D. and Nicol, C. J.(2008). Prevalence and factors associated with it, of birds dead on arrival at the slaughterhouse and other rejection conditions in broiler chickens. British Poultry Science, 49(6),685 – 696.

Igbokwe IO, Adawaren EO, Abba Y, 2012. Concurrent outbreak of staphylococcal pneumonia with infectious bursal disease in broiler chickens. Comparative Clinical Pathology, 21(6), 1571-1575.

Kalmar ID, Vanrompay D, Janssens GPJ, 2013. Broiler ascites syndrome: Collateral damage from efficient feed to meat conversion. Veterinary Journal. Article in Press.

Karaer Z, Guven E, Akcay A, Kar S, Nalbantoglu S, Cakmak A, 2012. Prevalence of subclinical coccidiosis in broiler farms in Turkey. Tropical Animal Health and Production, 44(3), 589-594.

Knowles TG, Kestin SC, Haslam SM, Brown SN, Green LE, Butterworth A, Pope SJ, Pfeiffer D, Nicol CJ, 2008. Leg disorders in broiler chickens: Prevalence, risk factors and prevention. PLoS One, 3(2), art. No. e1545.

Löhren, U, 2012. Overview on current practices of poultry slaughtering and poultry meat inspection. Supporting Publications 2012:EN-298, 58 pp. Available online: www.efsa.europa.eu/publications

Lupo C, Chauvin C, Balaine L, Petetin I, Péraste J, Colin P, and Le Bouquin S, 2008. Postmortem condemnations of processed broiler chickens in Western France. Veterinary Record, 162(22), 709-713.

Lupo C, Bougeard S, Balaine L, Michel V, Petetin I, Coiln P, LeBouquin S.and Chauvin C, 2010. Risk factors for sanitary condemnation in broiler chickens and their relative impact: application of an original multiblock approach. Epidemiology and Infection, 138 (3), 364-375.

Michel V, Prampart E, Mirabito L, Allain V, Arnould C, Huonnic D, Le Bouquin S and Albaric O, 2012. Histologically-validated footpad dermatitis scoring system for use in chicken processing plants. British Poultry Science, 53(3), 275-281.

Santana ÂP, Murata LS, Freitas CG, Delphino MK and Pimentel CM, 2008. Causes of condemnation of carcasses from poultry in slaughterhouses located in state of Goiás, Brazil. Ciencia Rural, 38(9), 2587-2592.

Silva JAO, Simões GS, Rossa A, Oba A, Ida EI, and Shimokomaki M, 2011. Preslaughter transportation and shower management on broiler chicken dead on arrival (DOA) incidence. (Manejo pré-abate de transporte e banho sobre a incidência de mortalidade de frangos de corte). Semina: Ciencias Agrarias, 32(2), 795-800.

Stegeman A, Bouma A, Elbers ARW, De Jong MCM, Nodelijk G, De Klerk F, 2004. Avian influenza A Virus (H7N7) Epidemic in the Netherlands in 2003: course of the epidemic and effectiveness of control measures. The Journal of Infectious Diseases, 190, 2088-2095.

Timbermont L, Lanckriet A, Gholamiandehkordi AR, Pasmans F, Martel A, Haesebrouck F, Ducatelle R, Van Immerseel F, 2009. Origin of Clostridium perfringens isolates determines the ability to induce necrotic enteritis in broilers. Comparative Immunology, Microbiology and Infectious Diseases, 32(6), 503-512.

Trampel DW, Smith SR, Rocke TE, 2005. Toxicoinfectious botulism in commercial caponized chickens. Avian Diseases, 49(2), 301-303.

van Breukelen, G. (2013). Sample Size Calculation for Prevalence Estimation in Three Level Population. In preparation.

van den Berg T, Houdart P, 2008. Avian influenza outbreak management: action at time of confirmation, depopulation and disposal methods; the "Belgian experience" during the H7N7 highly pathogenic avian influenza epidemic in 2003. Zoonoses Public Health, 55(1), 54-64.

Won GY, Moon BM, Oh IG, Matsuda K, Chaudhari AA, Hur J, Eo SK, Yu IJ, Lee YJ, Lee YS, Kim BS, Lee JH, 2009. Profiles of virulence-associated genes of avian pathogenic Escherichia coli isolates from chickens with colibacillosis. Journal of Poultry Science, 46(3), 260-266.

Xin H, Berry IL, Barton TL, Tabler GT, 1994. Feed and water consumption, growth, and mortality of male broilers. Poultry Science, 73(5), 610-616.

Yogaratnam V, 1995. Analysis of the causes of high rates of carcasses rejection at a poultry processing plant. The Veterinary Record, 137(9), 215-217.

APPENDICES

Appendix A. Commission Draft Document "Data collection in broilers' slaughterhouses- Implementation of Article 6.2 of Council Directive 2007/43/EC on the protection of chickens for meat production

DRAFT DOCUMENT

DATA COLLECTION IN BROILERS' SLAUGHTERHOUSES

IMPLEMENTATION OF ARTICLE 6.2 OF COUNCIL DIRECTIVE 2007/43/EC ON THE PROTECTION OF CHICKENS FOR MEAT PRODUCTION

1. Objectives of the data collection

To achieve an overall picture in the EU and each of the 27 Member States on the welfare of broilers over a period of 12 months.

This data collection is meant to be a onetime exercise, in which only criteria of on farm welfare will be considered. For this exercise, the previous work already performed in some Member States will be taken into account. Member States are invited to send the relevant information to the Commission. This information could be shared with EFSA.

The results of the data collection will be useful to define the abnormal levels mentioned in the Annex III of Directive 2007/43/EC.

2. Sample of flocks

The sample will be stratified into two categories: Member States and maximum stocking densities.

2.1. Geographical situations according to Member States

Data will be collected in each of the 27 Member States. Member States are free to further stratify geographically according to regions or by administrative areas relevant for them.

2.2. Densities of the flocks

The actual stocking densities of flocks at the moment of departure to the slaughterhouse will be recorded.

2.3. Seasonality

Data should be collected at least once a month in slaughterhouses for a period of 12 consecutive months.

Data will be analysed according to the evolution throughout the year.

2.4. Slaughterhouses

Data should be collected in at least 5 different slaughterhouses (if available).

In the case where there are less than 5 slaughterhouses in the Member State, all slaughterhouses killing broilers should participate in the data collection.

2.5. Number of flocks[8]

A flock is defined in Article 2j) of Directive 2007/43/EC as a group of chickens which are placed in a house of a holding and are present in this house at the same time. The objective is to detect and assess prevalence of conditions at EU level which occur with approximately 1% (or greater) prevalence within the flocks.

1% randomly selected flocks of the total number of flocks slaughtered each year should be tested with a minimum of 200 flocks (in the case where there are less than 200 flocks slaughtered each year in the

[8] EFSA will be asked to provide technical assistance to the Commission on the feasibility of achieving the objectives of the data collection to detect and assess prevalence of conditions at EU level which occur with approximately 1% (or greater when a percentage of 1% of the total number of flocks slaughtered each year in each MS, with a minimum of 200 flocks is sampled.

Member State, all flocks of this Member State should be tested) and 500 flocks maximum (in the case where 1% of the flocks represent more than 500 flocks).

2.6. Number of birds per flock

In order to have a sufficient precision:

For hock burns, 300[9] birds per flock will be assessed. Assessment can be made on different birds for each criterion (i.e. all the assessments do not need to be carried out on the same 300 birds).

For food pad dermatitis, 100 randomly selected feet per flock will be assessed, being careful not to assess the two feet of a same bird.

In regard to the sampling method, the Food Business Operator will be asked to organize the work in order to sample half of the animals at the beginning of the slaughter chain and half at the end.

3. Food chain information

The following data must be provided by the owner or keeper for each of the flocks tested. This information must be recorded in the data collected by the slaughterhouse:

- Date of slaughter
- Number of birds of the flock sent for slaughter
- Age of the birds of the flock at slaughter
- Hybrid or breed of the birds of the flock
- Country of origin of the birds
- Are these animals part of thinning[10]?
- Number of times thinning has occurred to this flock prior to the current thinning?
- The actual stocking density of the flock at the moment of departure to the slaughterhouse.
- Holding number of the flock
- Flock identification
- Cumulative daily mortality of the flock[11]
- Culling rate of the flock[12]
- Standard mortality of the flock
- Live weight at slaughter of the birds of the flock

4. Welfare indicators to be collected per flock

The indicators to be collected have been selected on the basis of the following criteria:

- The feasibility of collecting such indicators in slaughterhouses, on the basis of what has already been done in several Member States (UK, SE, ES, FR, DK, PT).

- Impact, in terms of animal welfare, of the indicators for which data are to be collected according to the number of birds affected the severity of the condition, and the correlation with management practices.

Seven welfare indicators will be collected.

[9] The final number of animals will be based on scientific evidence and discussed with EFSA.

[10] Thinning is defined as a farming practice, different from depopulation, used to meet the consumers' demand or to decrease the stocking density.

[11] As defined in the Directive 2007/43/EC. Note that farmers are obliged to give this information in the documentation accompanying the flock only in the case of stocking densities higher than 33 kg/m^2.

[12] Culling rate defined as the number of broilers culled until the day they are sent to slaughter divided by the number of broilers initially present in the house. The number of broilers initially present includes the birds delivered for free by breeding companies.

Welfare indicators
1. Hock burns
2. Foot pad dermatitis
3. Total rejections
4. Culling rate
5. Cumulative daily mortality
6. Standard mortality
7. Dead on arrival at the slaughterhouse

5. Assessment of welfare indicators

For each of the indicators, the place where the indicator can be assessed, the assessment rules and the training needs and the staff responsible for measuring the data should be defined.

The collection of indicators such as hock burns, food pad dermatitis, total rejections, dead on arrival at the slaughterhouse is the responsibility of the staff of the slaughterhouse or of the national competent authorities.

The collection of data such as culling rate, cumulative daily mortality and standard mortality is the responsibility of the owners or keepers and must be registered by the administrative staff of the slaughterhouse or by the national competent authorities.

In all cases, official veterinarians in slaughterhouses will be responsible for the supervision of the collection of indicators.

In all cases, training for the staff responsible of the data collection is needed. In the case of farmers, keepers or owners the training is part of the requirements of Article 4 of Directive 2007/43/EC.

For the staff in slaughterhouses, a one-day training session will be organized for all Member States and related countries, with at least two participants from each MS. An updating half-day training course will be organized two months later for the same participants. For the assessment of hock burns and food pad dermatitis, standardized pictures[13] shall be used.

In the case of total rejections, no additional training is needed for the staff except on rules for recording the results.

5.1. Hock burns

Indicator	Hock burns
Place of collection of data	Slaughter line after scalding or after plucking
Assessment rules	Scores 0, 1 (absence or presence)
	Score 0, no lesions or very small Area affected: none or very small Lesion description: absence or very small (up to 2 mm) superficial lesions, slight discoloration in a limited area (up to 2 mm)
	Score 1, presence of lesions Area affected: Lesion description: presence of superficial lesions (more than 2 mm), discoloration in an area bigger than 2 mm, some brown to

[13] Pictures available in Welfare Quality Assessment Protocol on Broilers and also from studies in UK, Sweden and Spain.

	black discoloration of the hock

5.2. Foot pad dermatitis

Indicator	Food pad dermatitis
Place of collection of data	After scalding on slaughter line or from feet collected in a box from the line.
Assessment rules	Scores 0,1 and 2[14]
	Score 0: healthy foot Area affected: none or very small Lesion description - very small superficial lesions (1-2 mm) - slight discolouration in a limited area - mild hyperkeratosis - completely healed scar
	Score 1: foot with mild lesion Area affected: does not extend over entire plantar pad Lesion description: - substantial discoloration of the foot pad - superficial lesion, no ulceration - dark papillae with no ulceration
	Score 2: foot with severe lesion Area affected: greater surface of plantar pad usually affected, sometimes with lesions on toes. Lesion description: - deeper lesions with ulceration, sometimes haemorrhage - scabs of significant size - severely swollen pad

5.3. Total rejections

Indicator	Total rejections
Place of collection of data	Slaughterhouse - Post mortem
Assessment rules	Number of carcasses rejected per flock indicating poor welfare conditions, such as abnormal levels of contact dermatitis, parasitism and systemic illness in the house. This definition excludes carcasses rejected due to the slaughter process. Total rejections include: Ante mortem rejects (culls/runts) Dead on arrival Abnormal colour/fevered Emaciation Ascites/oedema Bruising/Fractures

[14] Assessment rules similar as those used by DEFRA.

	Joint lesions/Arthritis Cellulitis Dermatitis Tumours/nodules Respiratory disease (Airsacuulitis) Salpingitis Hepatitis Perihepatitis Other (eg jaundice, Orgegon)

5.4. Culling rate

Indicator	Culling rate
Place of collection of data	At farm level/ recording of data in farm and in slaughterhouse
Assessment rules	Culling rate per flock from the date the animals entered the house until they left it for slaughter Culling rate means the number of chickens culled, either for disease or because of other reasons from the time the birds are placed in the house until the day the flock is sent to slaughter, divided by the number of chickens initially present in the house multiplied by 100. The number of chickens initially present includes the chickens delivered for free by breeding companies

Particular caution will be needed if thresholds for culling rates are defined at a later stage as culling animals to avoid unnecessary pain is good management practice.

5.5. Cumulative daily mortality

Indicator	Cumulative daily mortality
Place of collection of data	At farm level/ recording of data in farm and in slaughterhouse
Assessment rules	As defined in article 2 k) and l) of Directive 2007/43/EC Daily mortality rate means the number of chickens which have died in a house on the same day including those that have been culled either for disease or because of other reasons divided by the number of chickens present in the house on that day, multiplied by 100. Cumulative daily mortality rate means the sum of daily mortality rates. The number of chickens initially present includes the chickens delivered for free by breeding companies

5.6. Standard mortality

Indicator	Standard mortality
Place of collection of data	At farm level/ recording of data in farm and in slaughterhouse

Assessment rules	Standard mortality means the number of chickens found dead (including those that have been culled either for disease or because of other reasons) from the time the birds are placed in the house until the flock is sent to slaughter, divided by the number of chickens initially present in the house multiplied by 100. The number of chickens initially present includes the chickens delivered for free by breeding companies

5.7. Dead on arrival

Indicator	Dead on arrival
Place of collection of data	Slaughter line before shackling
Assessment rules	Number of dead birds on arrival. Annex III of Directive 2007/43/EC point 1.2 the number of broilers dead on arrival shall be recorded under the supervision of the official veterinarian. The dead on arrival rate means the number of chickens in the flock arriving dead at the slaughterhouse after transport divided by the number of chickens transported, multiplied by 100.

6. Future use of the data

The data collected will be aggregated and analysed at EU level.

The analysed data could be used to define trigger levels at which actions will be required from official veterinarians or owners/keepers.

In order to be able to aggregate the data, Member States must use the same data recording system.

Member States must not have more than 5% of missing data values for each variable per flock and the whole data base.

The complete data base per flock should be available to the Commission.

Member States should send each month a complete summary of the data collected.

Appendix B. Three level formula for sample size calculation

A tailored three level formula was developed to allow calculation of, in various defined scenarios of prevalence and prevalence variations between slaughterhouses and flocks, the number of units to be sampled for one level, given a defined number for the other two levels.

Such a formula allows undertaking a practical approach where representative sample sizes for the different levels can be set, on the basis of defined variability parameters. In particular, the formula allows computing the sample size at one level, given chosen sample sizes at each of the other two levels. It is therefore possible to fix those two sample sizes for which there are more constraints and compute the remaining sample size.

The 3 equations are reported below and full explanation on their possible application is given in Section 2.3.2 of the EFSA methodological report.

Since the sampling variance in these formulas is a function of the sample sizes at three levels (n_3 is the number of slaughterhouses, n_2 is the number of flocks within slaughterhouses and n_1 is the number of broilers/feet within flocks), we can only compute the sample size at one level, given some chosen sample size at each of the other two levels. For example, if the number of slaughterhouses and the number of broilers/feet to be sampled within each flock are constrained to be 10 and 100 at most, then fill in $n_3 = 10$ and $n_1 = 100$ to calculate n_2, the number of flocks to be sampled for each sampled slaughterhouse.

Considering the binary indicator of interest (i.e. presence of hock burn/foot pad dermatitis with score 2) for slaughterhouse i, flock j and broiler/feet k, where $i = 1, \cdots, n_3$, $j = 1, \cdots, n_2$ and $k = 1, \cdots, n_1$. The probability of the presence of the indicator for y_{ijk} is assumed to depend on the specific slaughterhouse, flock and broiler/feet that was sampled from, and is as follows:

$$P(y_{ijk} = 1) \equiv \pi_{y_{ijk}} = \frac{1}{1 + e^{-(\beta_0^C + a_i + b_{ij})}}$$

where β_0^C represents the overall conditional probability assuming no specific effect from flock and slaughterhouse, a_i and b_{ij} are the specific slaughterhouse and flock effects on the probability of presence of the indicator, which are assumed to be normally distributed with mean 0 and variances σ_S^2 and σ_F^2, respectively.

Equation 1- Number of slaughterhouses to be sampled (n_3), given the number of flocks to be sampled per slaughterhouse (n_2,) and given the number of broilers/feet to be sampled per flock (n_1):

$$n_3 = \left(Z_{1-\beta} + Z_{1-\alpha}\right)^2 \times \left(\frac{\sigma_S^2 + \frac{\sigma_F^2}{n_2} + \frac{\sigma_e^2}{n_2 n_1}}{(\beta_{0A}^C - \beta_{00}^C)^2} \right) \quad (1)$$

Where $\sigma_e^2 = [\pi^C(1 - \pi^C)]^{-1}$ is the variance at the broilers/feet level, π^C is the conditional prevalence given $a = b = 0$ (meaning absence of a slaughterhouse and flock effect on the prevalence) and using the true β_{0A}^C to compute π^C, and $\Delta = \beta_{0A}^C - \beta_{00}^C$ is the effect size;

Equation 2- Number of flocks to be sampled per slaughterhouse (n_2), given the number of slaughterhouses sampled (n_3), and given the number of broilers/feet sampled per flock (n_1):

$$n_2 = \left(Z_{1-\beta} + Z_{1-\alpha}\right)^2 \times \left(\frac{\sigma_F^2 + \frac{\sigma_e^2}{n_1}}{n_3 \Delta^2 - \left(Z_{1-\beta} + Z_{1-\alpha}\right)^2 \sigma_S^2} \right) \quad (2)$$

Equation 3- Number of broilers/feet to be sampled per flock (n_1), given the number of slaughterhouses sampled (n_3), and given the number of flocks sampled per slaughterhouse (n_2):

$$n_1 = \left(Z_{1-\beta} + Z_{1-\alpha}\right)^2 \times \frac{\sigma_e^2}{n_3 n_2 \Delta^2 - \left(Z_{1-\beta}+Z_{1-\alpha}\right)^2 \left(n_2 \sigma_S^2 + \sigma_F^2\right)} \quad (3)$$

Note that equation (2) can only be used if the denominator $n_3\Delta^2 - \left(Z_{1-\beta} + Z_{1-\alpha}\right)^2 \sigma_S^2 > 0$. This requires n_3 to be larger than according to equation (1) with $\sigma_F^2 = 0$ and $\sigma_e^2 = 0$. Similarly, the denominator of equation (3) must be larger than 0, which requires n_3 to be larger than according to equation (1) with $\sigma_e^2 = 0$. Furthermore, if all slaughterhouses in the MS are sampled, then $\sigma_S^2 = 0$, because there is then no sampling error at the slaughterhouse level anymore. Therefore, n_3 is the total number of slaughterhouses in the MS and equation (2) can be used to compute the total number of flocks needed, $n_3 * n_2$.

In the cases where the number of slaughterhouses sampled is more than 10% of the total population (and therefore the population cannot be considered infinite), then it is necessary to apply a 'correction factor' to that part of the sampling error which is due to prevalence variation between slaughterhouses.

More specifically, in the sample size equations the outcome variance at the slaughterhouse level (σ_S^2) must be multiplied with the following correction factor:

$$\frac{n_3 - 10}{n_3 - 1} \quad (4)$$

where total n_3 is the total number of slaughterhouses in the MS. The value "10" is used within the correction factor equation instead of 20, purely to maintain a conservative approach to the sample size as 10 is the minimum number of slaughterhouses to be sampled per MS (minimum n_3) and thus gives the weakest downward correction of the outcome variation at the slaughterhouse level and thereby also of the sample size needed.

The correction factor for the 2 level model (i.e. if all slaughterhouses are sampled would be the above equation (4), but now replacing the total n_3 in the MS with the total number of flocks in the MS, and replacing the value 10 with the value 200 in view of the minimum number of flocks to be sampled of 200 (minimum $n_3 * n_2$).

The theoretical derivations of equations (1), (2) and (3) are based on MQL1 estimation which is biased and it is known that PQL2 estimation, which is fairly unbiased, requires larger samples, depending on the amount of outcome variation at the highest design level. Therefore, simulations were performed in order to quantify the correction needed to adjust the theoretical formulas accordingly.

It is shown (EFSA 2013) that for the two level model, an increase of the total number of flocks sampled ($n_3 * n_2$) of at least 30% is needed in order to account for the approximation used in the theoretical derivations. In the three level model, an increase of at least 10% of the number of slaughterhouses sampled (n_3) is needed in order to account for the approximation used in the theoretical derivations. The fact that the correction is stronger for the two-level model than for the three-level model is due to the fact that we assumed the prevalence variation to be bigger at the flock level (i.e. 2.5 or 4.2) than at the slaughterhouse level (i.e. 1), based on the pilot data from France and the UK.

Finally, the sample size needs to be corrected for the loss of power incurred by size variation between slaughterhouses and between flocks, since such size variation affects the definition of prevalence at MS level and thereby also its estimator and the sampling error of this estimator (van Breukelen, 2013).

For three-stage sampling where the sampling probability per slaughterhouse is proportional to its annual throughput (number of broilers/year) whereas the sample size (n_2, n_1) is the same for each sampled slaughterhouse, the correction depends on the coefficient of variation (CV) of flock size within slaughterhouses. An increase of the number of flocks per slaughterhouse (n_2) by 20% is sufficient if the CV is 0.45, which is almost the maximum CV if flock size is normally distributed.

For two-stage sampling where all slaughterhouses are sampled and within each slaughterhouse the number of flocks sampled is proportional to the annual throughput, the correction depends on the CV of flock size variation within slaughterhouses as well as on the CV of slaughterhouse size variation (in terms of broilers/year). If both CV are 0.30, or if one is 0.45 and the other is 0, then increasing the total number of flocks sampled by 20% is still sufficient. If both CV are 0.45, then the total number of flocks needs to be increased by 44%.

Appendix C. Sufficient sample sizes and proposed sample plans by MS

In this Appendix, a table is provided for each MS, showing the results of the sample size assessment according to the procedure used to carry out the analysis. Moreover, for those MS where full data were available, tables by countries showing specific sampling plans are also included.

The procedure applied for the sample size calculation for each MS is described in Section 3.2.2 and 3.2.3 of this report and it is summarised in Table 1 of Section 3.2.3.

It should be noted that, for countries with more than 20 slaughterhouses and where a three level model is applied, the number of slaughterhouses to be sampled depends on the combination number of flocks per slaughterhouse and number of broilers/feet per flock chosen. Among the different combinations, the one considered the most appropriate by EFSA is to sample 24 flocks per slaughterhouse with 300 broilers/feet per flock. For each MS where this approach is used, the recommended combination is the first one highlighted in green in each MS table.

The assessment and the sampling plans are based on the latest data provided by MS; in most cases they were based on 2012 data. For FR, the assessment was based on data from 2010.

The sample size assessment was carried out for the following 19 MS: AT, BE, CY, CZ, DE, DK, EE, ES, FI, FR, IE, LV, LT, NL, PL, RO, SE, SI and UK.

The assessment shows that:

i) for 15 MS (AT, CY, CZ, DE, DK, EE, FI, IE, LT, LV, NL, PL, RO, SE and SI), 1% of the total number of flocks slaughtered in one year time is not a sufficient sample to assess 1% prevalence of the given conditions;

ii) for 3 MS (BE, ES and UK), the 1% sample is sufficient depending on the combination number of flocks and broilers/feet chosen;

iii) for 1 MS (FR), when any combination is used, 1% of the total number of flocks slaughtered in one-year time is a sufficient sample to assess a prevalence of 1%.

Sampling plans are proposed for 14 MS: AT, BE, CY, CZ, DK, ES, FI, FR, IE, LV, LT, RO, SI and UK in the tables of this Appendix. In particular, the sampling plans for those countries were designed following two approaches:

i) for MS with more than 20 slaughterhouses and where a three level model was applied for the assessment (BE, CY, CZ, ES, FR, RO, UK), the sampling plan is proposed only for the recommended sampling combination of 24 flocks to be tested per each slaughterhouse and 300 broilers/feet to be tested per flock. The number of slaughterhouses to be sampled varies for each country. Such a plan consists of the list of slaughterhouses to be sampled, based on a weighted random selection of slaughterhouses depending on their throughputs;

ii) for MS with less than 20 slaughterhouses which are all to be sampled (AT, DK, FI, IE, LV, LT, SI), the sampling plan consists of the allocation of the numbers of flocks to be tested in each slaughterhouse, depending on the slaughterhouse's throughput.

For a MS with only one broiler slaughterhouse, the sampling plan was not designed as all flocks are clearly taken to the same slaughterhouse. For 4 MS, the necessary information at slaughterhouse level to design possible sampling plans was not available.

Table 2: Sufficient flocks sample size in Austria

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
360	NO

Table 3: Proposed sampling plan for Austria

Anonymous slaughterhouse identifier	No of flocks to sample in one year time distributed by slaughterhouse throughput
AF	146
AA	107
AS	16
AL	93

Table 4: Sufficient flock sample sizes for Belgium, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	18	432	YES
24	400	18	432	YES
24	500	18	432	YES
36	100	19	684	YES
36	200	17	612	YES
36	300	17	612	YES
36	400	16	576	YES
36	500	16	576	YES
48	100	18	864	NO
48	200	16	768	YES
48	300	16	768	YES
48	400	16	768	YES
48	500	15	720	YES
60	100	17	1020	NO
60	200	16	960	NO
60	300	15	900	NO
60	400	15	900	NO
60	500	15	900	NO

Table 5: Sampling plan for Belgium for the recommended combination highlighted in green-weighted random selection of slaughterhouses to sample depending on their throughputs

Slaughterhouse- anonymous identifier	No of flocks to sample in each selected slaughterhouse
Ee	24
S	24
R	24
L	24
Hh	24
Jj	24
P	24
T	24
Cc	24
V	24
C	24
Mm	24
A	24
Kk	24
O	24
Oo	24
E	24
Tt	24

Table 6: Sufficient flock sample sizes for Cyprus, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	16	384	NO
24	400	15	360	NO
24	500	15	360	NO
36	100	17	612	NO
36	200	15	540	NO
36	300	14	504	NO
36	400	14	504	NO
36	500	14	504	NO
48	100	15	720	NO
48	200	14	672	NO
48	300	13	624	NO
48	400	13	624	NO
48	500	13	624	NO
60	100	14	840	NO
60	200	13	780	NO
60	300	13	780	NO
60	400	13	780	NO
60	500	13	780	NO

Table 7: Sampling plan for Cyprus for the recommended combination highlighted in green-weighted random selection of slaughterhouses to sample depending on their throughputs

Slaughterhouse- anonymous identifier	No of flocks to sample in each selected slaughterhouse
N	24
V	24
U	24
H	24
A	24
B	24
G	24
P	24
R	24
L	24
E	24
D	24
J	24
S	24
M	24
W	24

Table 8: Sufficient flock sample sizes for Czech Republic, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	16	384	NO
24	400	16	384	NO
24	500	15	360	NO
36	100	17	612	NO
36	200	15	540	NO
36	300	14	504	NO
36	400	14	504	NO
36	500	14	504	NO
48	100	16	768	NO
48	200	14	672	NO
48	300	14	672	NO
48	400	13	624	NO
48	500	13	624	NO
60	100	15	900	NO
60	200	14	840	NO
60	300	13	780	NO
60	400	13	780	NO
60	500	13	780	NO

Table 9: Sampling plan for Czech Republic for the recommended combination highlighted in green- weighted random selection of slaughterhouses to sample depending on their throughputs

Slaughterhouse- anonymous identifier	No of flocks to sample in each selected slaughterhouse
1	24
2	24
3	24
4	24
5	24
6	24
7	24
8	24
9	24
10	24
11	24
12	24
13	24
14	24
15	24
16	24

Table 10: Sufficient flocks sample size in Denmark

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
360	NO

Table 11: Proposed sampling plan for Denmark

Slaughterhouse- anonymous identifier	No of flocks to sample in one year time distributed by slaughterhouse throughput
1	139
2	117
3	101
4	4
5	1
6	1

Table 12: Sufficient flocks sample size in Estonia

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
206	NO

Estonia has only one poultry slaughterhouse, where all samples will have to be taken.

Table 13: Sufficient flocks sample size in Finland

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
360	NO

Table 14: Proposed sampling plan for Finland

Slaughterhouse- anonymous identifier	No of flocks to sample in one year time distributed by slaughterhouse throughput
A	190
B	128
C	44

Table 15: Sufficient flock sample sizes for France, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	20	480	YES
24	400	20	480	YES
24	500	19	456	YES
36	100	21	756	YES
36	200	19	684	YES
36	300	19	684	YES
36	400	18	648	YES
36	500	18	648	YES
48	100	20	960	YES
48	200	18	864	YES
48	300	18	864	YES
48	400	18	864	YES
48	500	17	816	YES
60	100	19	1140	YES
60	200	18	1080	YES
60	300	17	1020	YES
60	400	17	1020	YES
60	500	17	1020	YES

Table 16: Sampling plan for France for the recommended combination highlighted in green-weighted random selection of slaughterhouses to sample depending on their throughputs

Slaughterhouse- anonymous identifier	No of flocks to sample in each selected slaughterhouse
1	24
2	24
3	24
4	24
5	24
6	24
7	24
8	24
9	24
10	24
11	24
12	24
13	24
14	24
15	24
16	24
17	24
18	24
19	24
20	24

Table 17: Sufficient flock sample sizes for Germany, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	17	408	NO
24	400	16	384	NO
24	500	16	384	NO
36	100	18	648	NO
36	200	16	576	NO
36	300	15	540	NO
36	400	15	540	NO
36	500	14	504	NO
48	100	16	768	NO
48	200	15	720	NO
48	300	14	672	NO
48	400	14	672	NO
48	500	14	672	NO
60	100	15	900	NO
60	200	14	840	NO
60	300	14	840	NO
60	400	13	780	NO
60	500	13	780	NO

Table 18: Sufficient flocks sample size in Ireland

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
222	NO

Table 19: Proposed sampling plan for Ireland

Slaughterhouse- anonymous identifier	No of flocks to sample in one year time distributed by slaughterhouse throughput
2	17
3	78
4	25
5	102

Table 20: Sufficient flocks sample size in Latvia

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
207	NO

Table 21: Proposed sampling plan for Latvia

Slaughterhouse- anonymous identifier	No of flocks to sample in one year time distributed by slaughterhouse throughput
A	122
B	86

Table 22: Sufficient flocks sample size in Lithuania

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
244	NO

Table 23: Proposed sampling plan for Lithuania

Slaughterhouse-anonymous identifier	No of flocks to sample in one year time distributed by slaughterhouse throughput
1	7
2	83
3	23
4	14
5	4
6	113
7	2

Table 24: Sufficient flocks sample size in The Netherlands

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
360	NO

It should be noted that for The Netherlands, the total number of flocks slaughtered in one year time is an estimation calculated from the annual throughputs of broiler slaughterhouses in the country, excluding only two small establishments for which data were not available at the time of the request.

Table 25: Sufficient flock sample sizes for Poland, considering different combinations of number of flocks/number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	21	504	NO
24	400	21	504	NO
24	500	20	480	NO
36	100	22	792	NO
36	200	20	720	NO
36	300	19	684	NO
36	400	19	684	NO
36	500	19	684	NO
48	100	21	1008	NO
48	200	19	912	NO
48	300	19	912	NO
48	400	18	864	NO
48	500	18	864	NO
60	100	20	1200	NO
60	200	18	1080	NO
60	300	18	1080	NO
60	400	18	1080	NO
60	500	18	1080	NO

It should be noted that for Poland the total number of broiler flocks slaughtered in one year was not available. The assessment was therefore based on the total number of flocks kept on the farm and provided to EFSA.

Table 26: Sufficient flock sample sizes for Romania, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	18	432	NO
24	400	17	408	NO
24	500	17	408	NO
36	100	19	684	NO
36	200	17	612	NO
36	300	16	576	NO
36	400	16	576	NO
36	500	15	540	NO
48	100	17	816	NO
48	200	16	768	NO
48	300	15	720	NO
48	400	15	720	NO
48	500	15	720	NO
60	100	16	960	NO
60	200	15	900	NO
60	300	15	900	NO
60	400	14	840	NO
60	500	14	840	NO

Table 27: Sampling plan for Romania for the recommended combination highlighted in green-weighted random selection of slaughterhouses to sample depending on their throughputs

Slaughterhouse- anonymous identifier	No of flocks to sample in each selected slaughterhouse
AB	24
T	24
R	24
P	24
AD	24
F	24
AF	24
D	24
S	24
B	24
A	24
O	24
L	24
AG	24
Y	24
H	24
Z	24
G	24

Table 28: Sufficient flocks sample size in Slovenia

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
360	NO

Table 29: Proposed sampling plan for Slovenia

Slaughterhouse- anonymous identifier	No of flocks to sample in one year time distributed by slaughterhouse throughput
AM	28
PI	81
PP	253

Table 30: Sufficient flock sample sizes for Spain, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	20	480	YES
24	400	20	480	YES
24	500	20	480	YES
36	100	21	756	YES
36	200	19	684	YES
36	300	19	684	YES
36	400	18	648	YES
36	500	18	648	YES
48	100	20	960	NO
48	200	18	864	YES
48	300	18	864	YES
48	400	18	864	YES
48	500	18	864	YES
60	100	19	1140	NO
60	200	18	1080	NO
60	300	18	1080	NO
60	400	17	1020	NO
60	500	17	1020	NO

Table 31: Sampling plan for Spain for the recommended combination highlighted in green-weighted random selection of slaughterhouses to sample depending on their throughputs

Slaughterhouse- anonymous identifier	No of flocks to sample in each selected slaughterhouse
1	24
2	24
3	24
4	24
5	24
6	24
7	24
8	24
9	24
10	24
11	24
12	24
13	24
14	24
15	24
16	24
17	24
18	24
19	24
20	24

Table 32: Sufficient flocks sample size in Sweden

Tot number of flocks to sample in one year time	Is 1% of flocks slaughtered in one year a sufficient sample size?
360	NO

Table 33: Sufficient flock sample sizes for UK, considering different combinations of number of flocks and number of broilers/feet per flock

Number of flocks to sample per slaughterhouse in one year	Number of broilers/feet to sample per flock	Number of slaughterhouses to sample	Total flocks to sample	Is 1% of flocks slaughtered in one year a sufficient sample size?
24	300	20	480	YES
24	400	19	456	YES
24	500	19	456	YES
36	100	21	756	NO
36	200	19	684	NO
36	300	18	648	NO
36	400	18	648	NO
36	500	18	648	NO
48	100	19	912	NO
48	200	18	864	NO
48	300	18	864	NO
48	400	17	816	NO
48	500	17	816	NO
60	100	19	1140	NO
60	200	17	1020	NO
60	300	17	1020	NO
60	400	17	1020	NO
60	500	17	1020	NO

Table 34: Sampling plan for UK for the recommended combination highlighted in green- weighted random selection of slaughterhouses to sample depending on their throughputs

Slaughterhouse- anonymous identifier	No of flocks to sample in each selected slaughterhouse
ZZ	24
AA	24
B	24
D	24
X	24
J	24
Y	24
D1	24
SS	24
W	24
V	24
C1	24
S	24
NNN	24
U	24
SSS	24
YY	24
XX	24
JJJ	24
CCC	24

Appendix D. Theoretical sample sizes to detect different marginal prevalences

Table 34 in this Appendix D shows how the sample size strictly depends on the level of prevalence to be detected as well as on the effect size chosen.

In particular, sample sizes were computed for different combinations of number of slaughterhouses, flocks per slaughterhouse and broilers/feet per flock; the sample size calculation was done to detect different levels of prevalence under the null hypothesis (1%, 5% and 10%), considering 2 level of relative effect sizes of 1.5 and 2 (relative to the prevalence under the null hypothesis of 1%, 5% and 10%), and given 2 different levels of flock variability (2.5 and 4.2).

This theoretical calculation applies to a scenario where the number of slaughterhouses in the country can be considered infinite and based on chosen parameters, as described in this report. The table is provided in order to give an indication of the differences in sample sizes depending on the levels of prevalence to be detected.

Table 34: Sample size combinations and number of slaughterhouses to sample, assuming different effect sizes and variability parameters at flock level, for an infinite population of slaughterhouses

Flocks per SH	Broilers / feet per Flock	Number of slaughterhouses to sample											
		$\sigma_F^2 = 2.5$						$\sigma_F^2 = 4.2$					
		0.01 vs 0.015	0.01 vs 0.02	0.05 vs 0.075	0.05 vs 0.10	0.10 vs 0.15	0.10 vs 0.20	0.01 vs 0.015	0.01 vs 0.02	0.05 vs 0.075	0.05 vs 0.10	0.10 vs 0.15	0.10 vs 0.20
10	100	30	10	21	7	18	6	33	10	19	7	17	6
10	200	26	9	21	7	18	6	27	9	19	7	16	6
10	300	25	9	21	7	18	6	25	8	19	7	16	6
10	400	25	8	21	7	18	6	24	8	19	7	16	6
10	500	24	8	21	7	18	6	23	8	19	7	16	6
20	100	24	8	19	7	16	6	24	8	16	6	14	5
20	200	22	8	19	7	16	6	21	7	16	6	14	5
20	300	22	8	19	7	16	6	20	7	16	6	14	5
20	400	21	8	19	7	16	6	19	7	16	6	14	5
20	500	21	7	19	7	16	6	19	7	16	6	14	5
30	100	22	8	18	6	16	5	21	7	15	5	13	5
30	200	21	7	18	6	16	5	19	6	15	5	13	5
30	300	21	7	18	6	16	5	18	6	15	5	13	5
30	400	20	7	18	6	16	5	18	6	15	5	13	5
30	500	20	7	18	6	16	5	17	6	15	5	13	5
40	100	21	7	18	6	15	5	19	6	15	5	13	5
40	200	20	7	18	6	15	5	18	6	15	5	13	5
40	300	20	7	18	6	15	5	17	6	15	5	13	5
40	400	20	7	18	6	15	5	17	6	15	5	13	5
40	500	20	7	18	6	15	5	17	6	15	5	13	5
50	100	21	7	17	6	15	5	18	6	15	5	13	4
50	200	20	7	17	6	15	5	17	6	14	5	13	4
50	300	20	7	17	6	15	5	17	6	14	5	13	4
50	400	20	7	17	6	15	5	17	6	14	5	13	4
50	500	19	7	17	6	15	5	16	6	14	5	13	4

Suggested citation: European Food Safety Authority, 2013. Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses. EFSA Journal 2013; 11(7):3299, 68 pp. doi:10.2903/j.efsa.2013.3299

Available online: www.efsa.europa.eu/efsajournal

Appendix E. Stepwise procedure for sample size calculation in data collection system of welfare indicators in EU broilers' slaughterhouses

Based on the methodology and practical considerations illustrated in this report as well as on the full methodology reported in the EFSA methodological report (EFSA, 2013), this section summarises the steps that could be undertaken for computing the sample sizes needed for the data collection system of animal welfare indicators in EU broiler slaughterhouses.

The sample size calculation, as illustrated below, depends on the epidemiological parameters defined in the objectives of the data collection and used for the calculations. This stepwise procedure could therefore be used by MS as a tool to estimate the sample sizes needed for the data collection system of welfare indicators in EU broilers' slaughterhouses, depending on the designed prevalence to be detected, the chosen power of the test and the level of confidence, the chosen effect size and the variability parameters at slaughterhouse and flock level

This procedure refers to equations and formulas contained in specific Sections of the EFSA methodological report (EFSA, 2013).

Step 1: Choose the prevalence value at MS level according to H_0 and the smallest prevalence value which is just high enough above the H_0 value to make the difference between the two values worthwile detecting. For instance, choose $\pi = 0.01$ according to H_0 and $\pi = 0.015$ as smallest higher value if the study must be powerful enough to discriminate between these two values;

Step 2: Choose realistic values for the variance components σ_F^2 and σ_S^2, respectively the prevalence variation between flocks and between slaughterhouses on the logodds scale. When in doubt about realistic values, move on and check the values in step 4;

Step 3: Using $\pi^M = \frac{e^{\beta_0^M}}{1+e^{\beta_0^M}}$, compute the marginal prevalence on the logodds scale (β_0^M) . Then, using eq. (13) of the EFSA methodological report (EFSA, 2013) compute the conditional prevalence on the logodds scale (β_0^C). Repeat this for each of the two prevalence values from step 1;

Step 4: compute the 95% range of flock-specific prevalences $\left(\beta_{0L}^C, \beta_{0U}^C\right)$ on the logodds scale as follows: $\beta_{0L}^C \pm 2\sqrt{\sigma_S^2 + \sigma_F^2}$, and then transform this interval back to the (0,1) scale using $\pi_L^C = \frac{e^{\beta_{0L}^C}}{1+e^{\beta_{0L}^C}}$ and $\pi_U^C = \frac{e^{\beta_{0U}^C}}{1+e^{\beta_{0U}^C}}$. Check if this range is plausible for each of the two prevalence values of step 1. If not, then go back to step 2;

Step 5: Choose the test power (e.g. 80% or 90%) and test size α (e.g. 5% or 1% one-tailed), and then compute the sample size at each level using equations (1) to (3) reported in Appendix B of this report (the three levels formula) by fixing two of the three values, n_3, n_2 and n_1 and then computing the third;

Step 6: Increase the sample size at the highest level at which prevalence varies with, based on the simulations results, at least 30% for two level populations, and at least 10% for three level

Suggested citation: European Food Safety Authority, 2013. Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses. EFSA Journal 2013; 11(7):3299, 68 pp. doi:10.2903/j.efsa.2013.3299

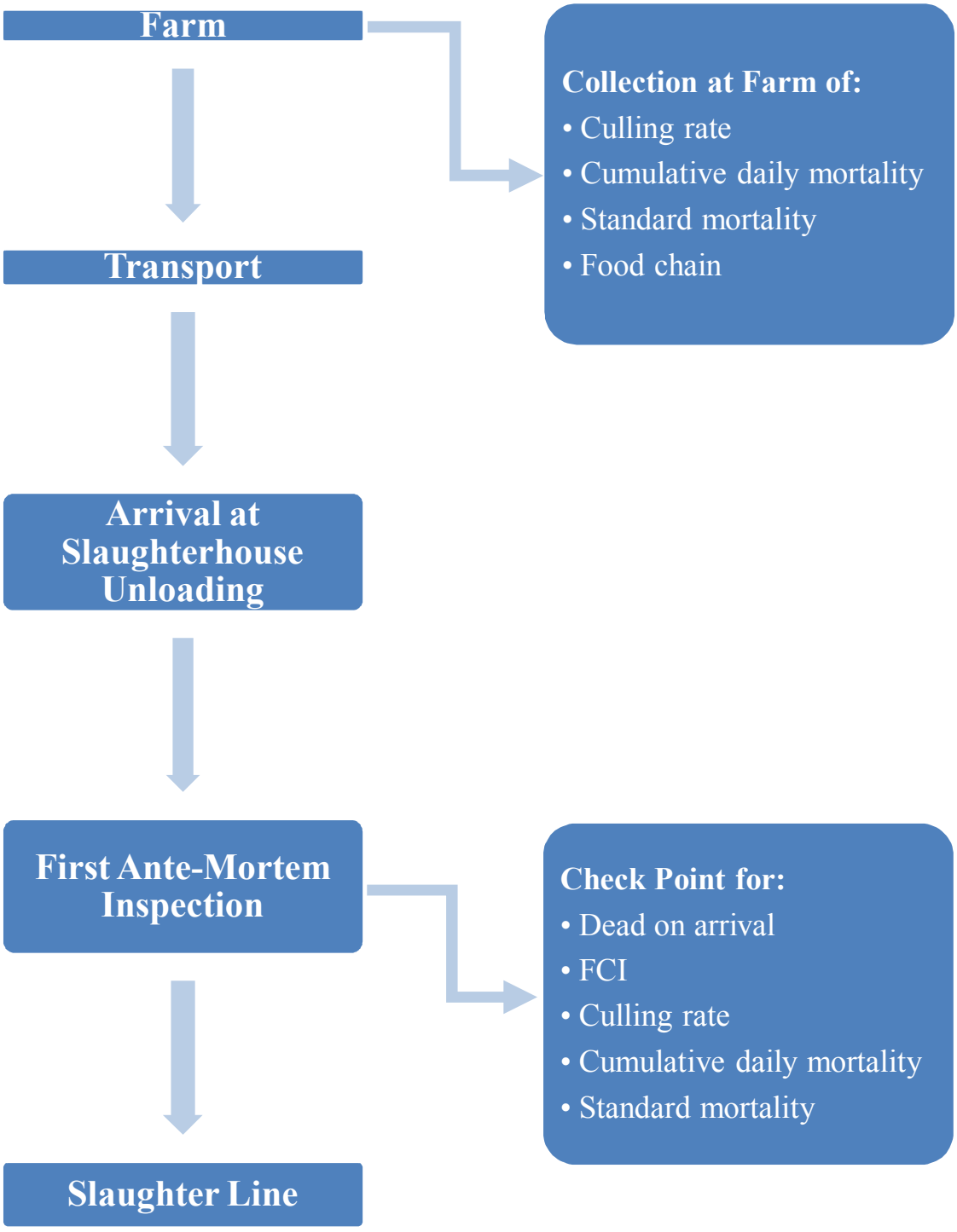
Available online: www.efsa.europa.eu/efsajournal

populations. So if $\sigma_S^2 > 0$ is assumed, then increase the number of slaughterhouses to be sampled (n_3) with 10%. If $\sigma_S^2 = 0$ is assumed (because all slaughterhouses are sampled), then increase the total number of flocks to be sampled ($n_3 * n_2$) by 30%.

Step 7: Finally, take into account size variation between slaughterhouses and between flocks as outlined at the end of Appendix B. As a rule of thumb this means increasing the number of flocks per slaughterhouse with 20% if a sample of slaughterhouses is taken, or increasing the total number of flocks by 20% up to 44% if all slaughterhouses are sampled.

Appendix F. Poultry slaughter-line flow

The speed of the line depends on the degree of the automation. In broiler slaughterhouses, it can be any speed, up to 13.000 broilers/hour (EFSA, 2012). Modified flow diagram from EFSA, 2012.



Suggested citation: European Food Safety Authority, 2013. Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses. EFSA Journal 2013; 11(7):3299, 68 pp. doi:10.2903/j.efsa.2013.3299

Available online: www.efsa.europa.eu/efsajournal


Slaughter Line
Shackling
Gas Stunning
Possible check point
Electrical Stunning
Shackling
Possible check point
Bleeding
Scalding and Plucking
First Post-Mortem Inspection
Transfer on the conveyor belt to clean area and feet removal
Recomended check point for foot pad dermatitis in boxes
Recomended check point for hock burn
Hanging by the hocks on the shackles of the evisceration line
Second Post-Mortem Inspection
Third Post-Mortem Inspection
Evisceration
Check point for Total rejections
Transport Line
Check Point for:
• Dead on arrival
• FCI
• Culling rate
• Cumulative daily mortality
• Standard mortality
Cooling

Appendix G. Proposed Data model to submit the data to EFSA through the EFSA Data Collection Framework

Flock

Element Name	Data Type	Controlled Terminology	description	mandatory
progLegalRef	String(250)	Council Directive 2007/43/EC of 28 June 2007 laying down minimum rules for the protection of chickens kept for meat production. OJ L182, 19-28.	Reference to the legislation for the survey	Y
sampCountry	String(2)	COUNTRY[15]	Country where the slaughterhouse is located	Y
slaughterName	String(250)		Name of the slaughterhouse	Y
slaughterHseID	String(50)		Unique identifier for slaughterhouse – must be unique for the reporting country	Y
cuttingPlant	String(1)	Y/N	The slaughterhouse includes also a cutting plant	
stunningType	String(50)	gas, electrical waterbath, gas and electric, not reported, other	Type of stunning system used	
slaughterRegion	String(5)	NUTS[16]	NUTS 2 region where slaughterhouse is located	Y
capacityBroilers	Integer(20)		Slaughterhouse capacity - Total throughput expressed in number of broiler flocks slaughtered within the last 12 months	Y
sampY	Integer(4)		Year of slaughter of the flock	Y
sampM	Integer(2)		Month of slaughter of the flock	Y
sampD	Integer(2)		Day of slaughter of the flock With very large sized flocks, birds may be transported to the slaughterhouse over a number of days (this would not represent thinning).	Y

[15] http://www.efsa.europa.eu/en/efsajournal/pub/1457.htm
[16] http://www.efsa.europa.eu/en/efsajournal/pub/1457.htm

Suggested citation: European Food Safety Authority, 2013. Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses. EFSA Journal 2013; 11(7):3299, 68 pp. doi:10.2903/j.efsa.2013.3299

Available online: www.efsa.europa.eu/efsajournal

			In such cases, a record should be created for each day when parts of the flocks are transported to slaughter.	
holdingID	String(50)		Official code for the holding of the flock	Y
holdingRegion	String(5)	NUTS	NUTS 2 region where holding is located	Y
flockID	String(50)		Identifier for each flock sampled – must be unique for each country and allow identification of flocks sample more than one in the year	Y
sampMatCode	String(5)	A0C78 = Gallus gallus broiler	Five character code for the species slaughtered	Y
breedCode	String(50)	Cornish cross / Cornish-Rock / Cobb 500 / Cobb 700 / Ross 308 / Ross 708 / Ross PM3 / Arbor Acres / Label Rouge / Certified chickens / Other slow-growing breed / Other	Code to identify hybrid or breed of broilers of the flock	Y
breedText	String(250)		If other, specify the hybrid or breed of the broilers of the flock	
origCountry	String(2)	COUNTRY	Country of origin of the broilers flock	Y
age	Integer(5)		Age of the broilers of the flock, at date of slaughter, in days	Y
thinning	String(1)	Y/N	Slaughtered birds are part of a thinning (Thinning is defined as a farming practice, different from depopulation, used to meet the consumers' demand or to decrease the stocking density). This data refers to the current thinning being performed on those animals arrived to the slaughterhouse.	Y
N_thinning	Integer(20)		If thinning = Y specify the number of times thinning has occurred in the flock prior to current thinning. This data refers to the previous thinning performed on the flock of origin of those animals arrived at the slaughterhouse	
stockingDensity	String(5)	Stocking Density (see table below)	Code to classify stocking density of flock at moment of departure to the slaughterhouse	Y
usableArea	Double(20,10)		Usable area in which the flock is living during their life-time (expressed in meters)	Y
N_broilers	Integer(20)		Number of broilers in the flock sent for slaughter	Y
standardMortality	Double(20,10)		Number of dead chickens (including those that have been culled either for disease or because of other reasons) from the time the birds are placed in the house until the flock is sent to slaughter, expressed as a percentage of the number of chickens	Y

			initially present in the house (including chickens delivered for free by breeding companies)	
cullingRate	Double(20,10)		Number of chickens culled, either for disease or because of other reasons from the time the birds are placed in the house until the flock is sent to slaughter, expressed as a percentage of the number of chickens initially present in the house (including chickens delivered for free by breeding companies)	Y
flockWeight	Double(20,10)		Live weight at slaughter of the birds of the of flock (expressed in kilograms)	Y
N_DOA	Integer(20)		Number of broilers in the flock dead on arrival at slaughterhouse, expressed as a percentage of the number of broilers transported	Y
N_rejects	Integer(20)		Number of whole carcasses rejected	Y
contactDermatitis	Integer(20)		Number of carcasses rejected for reasons of contact dermatitis	
systemicIllness	Integer(20)		Number of carcasses rejected for reasons of systemic illness (septicaemia, airsacculitis, hepatitis, pericarditis, abscesses, tumours)	
parasitism	Integer(20)		Number of carcasses rejected for reasons of parasitism	
pointOfSampleHB	String(50)	postevisceration, postchilling, ante-mortem , post-mortem, postscalding, post-plucking, postwashing, prescalding, predefeathering, preevisceration, prechilling	Point in the process where carcass was sampled for hock burn	Y
pointOfSampleFP	String(50)	postevisceration, postchilling, ante-mortem , post-mortem, postscalding, post-plucking, postwashing, prescalding, predefeathering, preevisceration, prechilling	Point in the process where carcass was sampled for foot pad dermatitis	Y
prHockBurn	Integer(20)		Number of broilers sampled with observations of presence of hock burn (presence of superficial lesions (more than 2 mm),	Y

			discoloration in an area bigger than 2 mm, some brown to black discoloration of the hock)	
abHockBurn	Integer(20)		Number of broilers sampled with observations of absence of hock burn (absence or very small (up to 2 mm) superficial lesions, slight discoloration in a limited area (up to 2 mm))	Y
footpadZero	Integer(20)		Number of feet sampled with foot pad dermatitis score 0 (- very small superficial lesions (1-2 mm), slight discolouration in a limited area, mild hyperkeratosis, completely healed scar)	Y
footpadOne	Integer(20)		Number of feet sampled with foot pad dermatitis score 1 (- substantial discoloration of the foot pad, superficial lesion, no ulceration, dark papillae with no ulceration)	Y
footpadTwo	Integer(20)		Number of feet sampled with foot pad dermatitis score 2 (- deeper lesions with ulceration, sometimes haemorrhage, scabs of significant size, severely swollen pad)	Y

Stocking density codes

code	description
STD1A	≤ 33 kg/m^2
STD2A	33 kg/m^2 < x $\leq$ 36 kg/m^2
STD3A	36 kg/m^2 < x $\leq$ 39 kg/m^2
STD4A	[39] g/m^2 < x $\leq$ 42 kg/m^2

Appendix H. Correlation between cumulative daily mortality and standard mortality

This Appendix explains the statistical interpretation and the correlation between cumulative daily mortality and standard mortality. In particular, it provides:

Cumulative daily mortality:
"Daily mortality" is defined as "the number of chickens that have died in a house on the same day (including culling) divided by the number of chickens present in the house on that day, multiplied by 100". In biostatistics this is known as the hazard rate *h(t)* at day *t*. "Cumulative daily mortality" is defined as "the sum of the daily mortality rates", here to be denoted as *H(t)* for the cumulative daily mortality at day *t,* since placement in the house.

Standard mortality:
"Standard mortality" is defined as "the number of chickens found dead (including culling) from the time the birds are placed in the house until the flock is sent to slaughter, divided by the number of chickens initially present in the house multiplied by 100". Ignoring the final multiplication by 100, which only serves to transform a proportion into a percentage, this standard mortality equals *F(t) = 1-S(t)*, where *S(t)* is the survivor function at day *t* , that is, the probability of still being alive at day *t*.

Given the absence of censoring (i.e. given that all broilers are followed up until death), the two indicators, cumulative daily mortality *H(t)* and standard mortality *F(t)* , come down to the same thing when measured at the same time point, since their mathematical relation is as follows:

$$F(t) = 1 - S(t) = 1 - e^{-H(t)},$$

where:
$H(t) = \int_0^t h(x)\,dx$ if time is measured continuously, or:
$H(t) = \sum_{x=1}^{t} h(x)$ if time is measured in days

Table 35: Illustrative example considering a hypothetical small flock to show the relation between both indicators

Day	Flock size	No of deaths	Daily mortality	Cumulative daily mortality
1	500	10	0.0200	0.0200
2	490	14	0.0286	0.0486
3	476	8	0.0168	0.0654
4	468	12	0.0256	0.0910
5	456	9	0.0197	0.1108
6	447	7	0.0157	0.1264
7	440	11	0.0250	0.1514
8	429	18	0.0420	0.1934
9	411	4	0.0097	0.2031
10	407	7	0.0172	0.2203
	400	100		

Suggested citation: European Food Safety Authority, 2013. Technical assistance to the Commission (Article 31 of Regulation (EC) No 178/2002) for the preparation of a data collection system of welfare indicators in EU broilers' slaughterhouses. EFSA Journal 2013; 11(7):3299, 68 pp. doi:10.2903/j.efsa.2013.3299

Available online: www.efsa.europa.eu/efsajournal

The cumulative daily mortality at t =10 is here $H(t) = 0.2203$, which gives as standard mortality: $F(t) = 1 - e^{-0.2203} = 0.198$. Multiplied by 100 this gives 19.8%.
To compare, the standard mortality at t = 10 defined as the total number of deaths divided by the initial flock size, is 100/500 = 0.20, and multiplied by 100 this gives 20%.

The minor deviation between these two results, the mortality $F(t)$ as computed from the cumulative daily mortality $H(t)$, and the standard mortality as computed from the total percentage of deaths between placement in the house and transport to slaughter, is due to the fact that the equation assumes continuous time whereas the example uses time in days running from 1 to 10, which is discrete.

Incidentally, in this example the cumulative daily mortality $H(t)$ is itself already very close to the standard mortality even without transformation, that is 0.22 versus 0.20. A plot of $F(t)$ versus $H(t)$ for $H(t)$ from 0 to 1 in steps of 0.01 shows that $F(t) \approx H(t)$ as long as $H(t) < 0.10$.

Therefore, if the cumulative daily mortality is below 10%, then standard mortality and cumulative daily mortality are approximately the same. The cumulative daily mortality and the standard mortality are always related to one another, following the relationship $F(t) = 1 - e^{-H(t)}$.

Given the precise and simple relation between the cumulative daily mortality and the standard mortality (due to the absence of censoring), the data collection could be limited to the one indicator which is most easily measured, being most probably the standard mortality.

However, if interest is not only in the cumulative mortality on the day of slaughter, but also on the whole time course of the mortality from the first day the flock is placed in the house till the day of slaughter, e.g. to distinguish between early and late mortality, then the total series of cumulative daily mortality from day 1 till slaughter is needed.

After all, the standard mortality is recorded at the day of slaughter only. In terms of the table above, this means that the complete last column would be needed, not just the last number in it.