PEER REVIEWED

# Study Design & Data Analysis for EFSA Feed Additive Applications

The design of target animal safety and efficacy studies for EFSA feed additive applications requires the use of rigorous scientific methodology as well as the application of EFSA guidance documents that specify the study duration per species, endpoints according to the feed additive functional group/claims, and statistical analysis, among others. When an applicant fails to comply with the guidelines, without sound scientific justification, this can lead to a stop of the whole EFSA evaluation procedure whereby EFSA requests additional information (also known as a clock-stop). These studies are expensive and require a considerable amount of investment in terms of time from planning through execution and final data collection; thus getting the study design wrong can lead to a significant loss of time and money, and could also prevent the applicant from getting the authorisation to enter the EU market.

Feed additives in the EU can exert an effect on feeding stuffs or in animals. For the latter, it is important to show that the compounds for which an authorisation is sought are safe and efficacious for the target species, whether livestock or pets. This requires careful design of studies considering the requirements of the recently updated EFSA safety and efficacy guidelines[1,2]. Among others, the guidelines describe details for:

- N° and type of treatments,
- Study duration according to target species,
- Physiological stage of animals to use,
- Appropriate endpoints for each functional group,
- N° of studies and geographical location required.

**Target Animal Safety Studies**

Applicants designing target animal safety (TAS) studies must include a minimum of three treatment groups: control (without additive), maximum dose (use-level) and tolerance dose (x10-x100). The choice of tolerance dose will depend on the nature of the active substance (e.g. microorganisms, chemicals, minerals…) which could be determined using data from pilot in-house studies and/or scientific literature. Among species, the minimum study duration is for pets (28 days) while the maximum is for sows (approx. 140 days). For these trials, EFSA recommends using the most sensitive physiological stages for each species (e.g. cattle = weaned male bovines, chickens for fattening = only male broilers…), whereas the type of endpoints will depend on the tolerance dose chosen (see Table 1). Applicants are required to perform only one safety study per target animal species, and results can be extrapolated to physiologically related groups (see Figure 1).

| Tolerance dose in relation to the use-level | Type of endpoints |
|---|---|
| ≥100 | Check of clinical signs and record of mortality and performance parameters (body weight, feed intake, feed conversion ratio, average daily gain…) |
| 10 to <100 | Same as above + haematology and blood biochemistry |
| ≤10 | Same as above + gross pathology and histopathology |

*Table 1. Type of endpoints for a safety trial according to the tolerance dose tested*

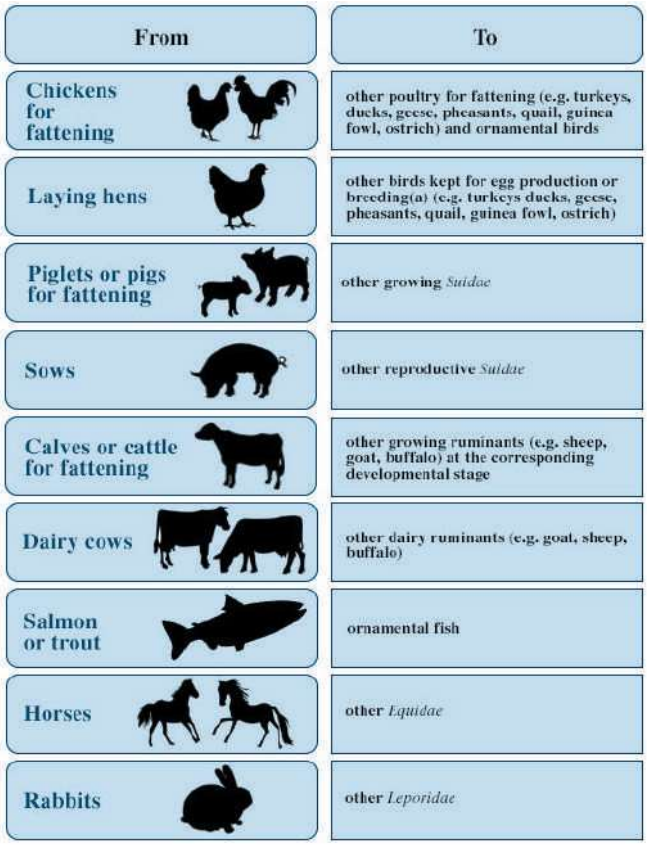



*Figure 1. Extrapolation allowed in safety and efficacy studies*

**Efficacy Studies**

On the other hand, a study aimed at showing efficacy of a feed additive will need to consider the inclusion of at least two treatments: control (without additive) + minimum effective dose (use-level). The study duration may not be the same as for safety studies (e.g. pigs for fattening (a) safety = 42 days; (b) efficacy = not less than 70 days) while, as a positive note for some applicants, there is the possibility to choose between long- and short-term studies depending on the functional group of interest. The physiological stage of animals required in efficacy trials will be mainly determined by age, and extrapolation of results to related species is also allowed. Regarding endpoints, the applicant will need to select those specific for the additive category and functional group, as well as some depending on the type of claims (e.g. for zootechnical additives with effects on product quality/composition = nutrient content, water-binding capacity, meat taste…). For the efficacy assessment, EFSA requires a minimum of three studies with significant results that should be performed in at least two different locations, one of which should be in the EU.

**Experimental Design: Experimental Units and Sample Size Calculations**

The information about the type of trial (e.g. TAS or efficacy), target animal category and endpoints that will be measured will inform the overall experimental design, experimental unit and the methodology required

for sample size calculations and statistical analysis. Firstly, applicants and researchers directing trials on-site should be careful when defining the experimental unit (Figure 2).

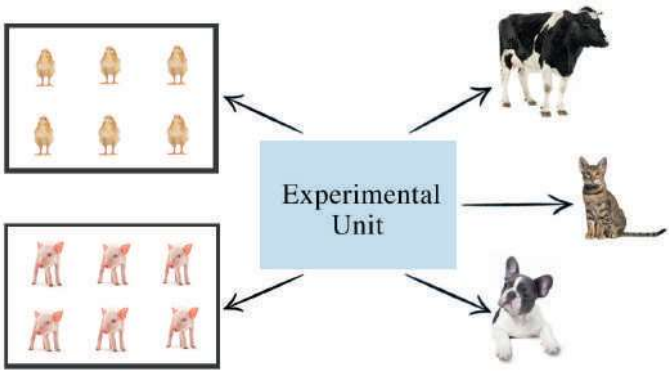



*Figure 2. Graphical representation of how an experimental unit is usually defined depending on the target animal species.*

For EFSA, if animals are penned in groups and all the animals in the pen share the same feed source (and feed intake is not measured individually), then the experimental unit for all parameters is the pen, not the individual animal[2]. Thus, in practical terms this means that, example, body weight measured individually throughout a trial will always be analysed on a pen basis if feed intake cannot be recorded per animal. In general, the consideration of individual replicates for statistical purposes is possible for studies performed in dogs, cats and dairy cows due to economic and practical reasons.

The experimental unit in a trial is also known as replicate. In order to estimate the minimum number of replicates necessary to detect relevant differences between treatments, applicants should perform sample size calculations that will be submitted to EFSA in the final study report. There are two types of formulas for these calculations depending on the scientific hypothesis to test, which could be (a) equivalence or non-inferiority or (b) superiority. Safety trials are aimed at testing if animals in the tolerance group have the same performance (equivalence) or do not perform worse (non-inferiority) than those in the control group, whereas efficacy trials will test if animals fed the test item (additive) perform better (superiority) than the control group. The formulas for both cases (see Table 2) have three things in common that the applicant needs to know:

- The standard deviation of the parameter(σ),
- The significance level (α),
- The statistical power (β).

In addition,

- For safety: the largest difference that is clinically acceptable between means, so that a difference bigger than this would matter in practice (d),
- For efficacy: the smallest difference between means that the applicant wishes to detect (δ).

| Hypothesis | Type of trial | Formula for sample size calculation[3] |
|---|---|---|
| Equivalence or non-inferiority[1] | Safety | $n = f(\alpha, \beta) \times 2 \times \sigma^2 / d^2$ |
| Superiority[2] | Efficacy | $n = \frac{2\sigma^2(z_\alpha + z_\beta)^2}{\delta^2} = \frac{2\sigma^2(1.96+0.84)^2}{\delta^2} = 15.7\frac{\sigma^2}{\delta^2}$ |

[1]Julious, S.A (2004); [2]Snedecor & Cochran (1967); [3]Symbol description: $n$ = sample number, σ = standard deviation, δ = smallest difference between means, α = 0.95 (significance level for majority of species), β = 0.80 (statistical power for majority of species), $z_{0.95}$ & $z_{0.8}$ are constants from a normal table, d = non-inferiority limit, $f(\alpha, \beta) = [\Phi^{-1}(\alpha) + \Phi^{-1}(\beta)]^2$, $\Phi^{-1}$ is the cumulative distribution function of a standardised normal deviate.

*Table 2. Formulas for sample size calculations according to the hypothesis tested*

The standard deviation of the parameters or endpoints that will be measured can be estimated from studies already performed using the same additive, scientific literature or previous similar studies performed at the experimental facility. The significance level (α) and statistical power (β) are related to the Type I and Type II errors respectively. Thus, α is the probability of rejecting the null hypothesis when it is true, whereas β is the probability of accepting the null hypothesis when it is false. For animal research is generally accepted the assumption of α = 0.95 and β = 0.80[3] (i.e. statistical significance is accepted at P≤0.05), which also applies for trials that will be submitted for EFSA assessment. However, for ruminants, pets and non-food producing animals, given that the phenotypical variability for these target animals is assumed to be larger compared to poultry and swine, EFSA accepts calculations assuming α = 0.90 and β = 0.75; (i.e. statistical significance is accepted at P≤0.10). This methodology for sample size calculation applies only to continuous variables and not to proportions and other types of data [4].

Once the total number of replicates has been estimated, animals will need to be allocated to the experimental treatments. A complete randomised block design is preferred, where the study director must ensure the utilisation of proper methods to randomise animals to pens (if applicable) and treatments. Before randomising the experimental units, animals can be stratified by body weight, gender, parity... if necessary to ensure homogeneity among pens and treatment groups. It is recommended to blind the team in charge of monitoring the study to the treatment information using, for example, colour codes to reduce bias.

**Handling Raw Data**

After the trial finishes and the relevant information is collected, it is important to make sure that all raw data has been recorded properly and in line with the study protocol. The files must contain clear names, units of measurement, a glossary to allow easy identification of endpoints, age at each measurement, treatment groups, pen numbers, reasons for mortality... For parameters calculated using the raw data (e.g. average daily gain, feed conversion ratio...), the dataset should include the formulas used. Failing to do this will affect the accuracy and interpretation of results, as well as the EFSA assessment, as EFSA will receive and check all supporting files.

**Statistical Analysis**

With a clean and accurate dataset, the applicant can proceed to perform statistical analyses. Depending on the type of variables measured (continuous, binomial, nominal...) the applicant will need to choose an appropriate hypothesis test (t-test, chi-square, ANOVA...) to determine if there are relevant and significant differences between treatments. The ANOVA (analysis of variance) is commonly used to assess the variability between and within treatments for fixed effects models (e.g. consideration of the effects of sex, month-season, genotype... on the dependent variable), and additionally REML is recommended for general linear mixed models (i.e. fixed + random effects such as litter, block...). Posteriorly, if results require the establishment of the pattern of differences between means, EFSA requires the performance of multiple pairwise comparisons with tests such as Scheffé and Tukey, while for comparisons with the control group the applicant could use Dunnet or Scheffé. Duncan & Fisher's Least Significant Difference are not recommended. Procedures such as the removal of outliers as well as data transformation to avoid the use

of non-parametric methods (e.g. log transformation for microbiological data) should always be described in the final report providing proper justification. The applicant must not forget that the statistical methodology has to be stated *a priori* in the trial protocol, and that all statistical printouts will need to be made available to EFSA.

**Pooling Data**

Even after careful planning, the statistical analyses can reveal that no significant differences were found between treatments during the trial. For these cases there is an alternative: the applicant can pool data from two or more non-significant studies as long as they have a similar design in terms of number of treatments (specific dose), target animal species, study duration, and endpoints measured. If there are no interactions between treatment and study, and the pooled data shows significant differences among treatments for relevant parameters, the results of the "pooled study" can be submitted to EFSA as one of the three individual trials required for the efficacy assessment of a feed additive application.

**Clock-stops & Getting it Wrong**

A key component of the EFSA evaluation procedure consists of checking if all the previously described requirements were addressed by the applicant. If they were not, or if some procedures were described vaguely or supporting evidence was not submitted properly (e.g. raw data and statistical printouts), the expert panel will stop their assessment and request that the applicant submits a clarification or additional information to support the study. This pause is called a clock-stop and could last weeks or years, depending on the complexity of the data/information requested. Below there are some real-life examples of what a clock-stop request can look like:

*"...the experimental unit was wrongly assigned to the individual animal... the applicant is invited to reanalyse the data using the correct unit, i.e. the pen..."*

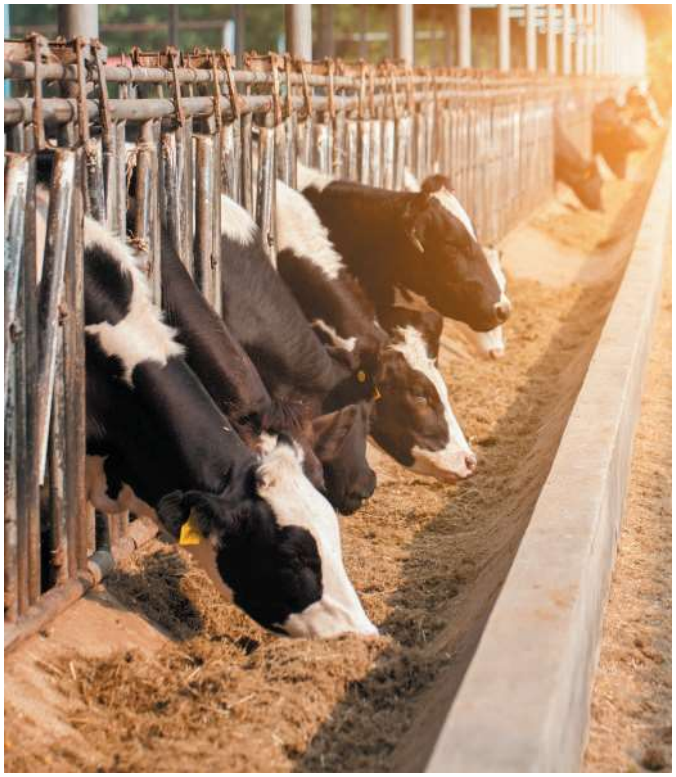
*"...provide clarifications on the statistical analysis performed including clarifications on the experimental unit and the test used to separate the means...finally, please confirm that no animals died during the study..."*

*"According to the report hens were randomly distributed between treatment groups but no information was given on how this randomisation was done. Clarify how the animals were distributed to the treatments and provide data on laying hen intensity and egg weight at study start."*

As shown above, an EFSA-compliant animal study for a feed additive dossier needs thoughtful consideration of study design, randomisation, raw data processing and collection, as well as the method of statistical analyses. EFSA guidelines for the assessment of safety and efficacy are some of the most valuable resources that the applicant will need to use to ensure compliance with the authorities' requirements and expedite entering the EU market.

**REFERENCES**


1. EFSA. Guidance on the assessment of the safety of feed additives for the target species. EFSA Journal 2017;15(10):5021 (2017)
2. EFSA. Guidance on the assessment of the efficacy of feed additives. EFSA Journal 2018;16(5):5274 (2018)
3. Blasco, A. Análisis de datos experimentales para proyectos fin de carrera. Editorial UPV, España. Ref: 473 (2012)
4. Whitley, E. and Ball, J. Statistics review 4: Sample size calculations. Critical Care Vol 6, N° 4, 335-341(2002)
5. Snedecor G.W. and Cochran W.G. Statistical Methods. Iowa State University Press, 6th ed. (1967)
6. Julious, S.A. Sample sizes for clinical trials with Normal data. Statist. Med, 23:1921-1986 (2004)


## Dr Eliana Henriquez

An Agricultural Engineer with applied experience in animal production and research. She holds an MSc in Animal Breeding and Reproduction Biotechnology as well as a PhD in Agricultural and Food Science and Technology. Eliana has worked for the pig industry in Venezuela, and trained with animal breeding and agricultural research organisations in the USA and Ireland. In her role at Pen & Tec Consulting, Eliana is responsible for providing regulatory advice to clients on study design and statistics, and manages feed additive registration projects.

**Email: eliana@pentec-consulting.eu**

## Dr Hannah Lester

A research scientist with over 10 years' experience in running veterinary clinical trials for regulatory approval. She has a PhD in parasite epidemiology and an MSc in Veterinary Public Health. Hannah has managed contract research organisations in the UK, and has also worked for government, providing veterinary policy advice. In her role as Scientific Director at Pen & Tec Consulting, Hannah is responsible for providing strategic advice on study design and statistics and manages feed additive and novel food registrations in the EU.

**Email: hannah@pentec-consulting.eu**