European Food Safety Authority

EFSA Journal 2011;9(12):2438

SCIENTIFIC OPINION

Guidance on conducting repeated-dose 90-day oral toxicity study in rodents on whole food/feed [1]

EFSA Scientific Committee[2, 3]

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

Following a request from the European Commission, the Scientific Committee developed principles and guidance for the establishment of protocols for 90-day whole food/feed studies in rodents. The guidance complements the OECD Test guideline 408 and provides specific advice for performing and reporting experiments carried out with whole food/feed. The conduct of the study and documentation should follow good laboratory practice. Preparation of appropriate test diets is a key element of the experiment with respect to the choice of the diet type, nutritional balance and necessary adjustments, processing, and storage. Since it is often not possible to include whole foods in high amounts in the test diet, fewer dose levels but more animals in control and top dose groups should be used to maximise the power of the study. In accordance with the European Directive 2010/063, the test animals should be housed socially. The guidance recommends that animals of the same sex are housed in pairs. A comprehensive set of endpoints as set out in the OECD TG 408 should be measured during and at the end of the 90-day period, as appropriate. The use of a power analysis to estimate a sample size capable of detecting a pre-specified biologically relevant effect size with a specified power and significance level is described. A randomised block design is recommended for 90-day toxicity studies when testing whole food/feed.

KEY WORDS

Repeated-dose toxicity study, 90-day, rodent, whole food/feed, experimental design, statistical analysis

[1] On request from the European Commission, Question No EFSA-Q-2009-00941, adopted on 9 November 2011.

[2] Scientific Committee members: Boris Antunović, Sue Barlow, Andrew Chesson, Albert Flynn, Anthony Hardy, Michael Jeger, Ada Knaap, Harry Kuiper, David Lovell, Birgit Nørrung, Iona Pratt, Ivonne Rietjens, Josef Schlatter, Vittorio Silano, Frans Smulders and Philippe Vannier. Correspondence: scientific.committee@efsa.europa.eu

[3] Acknowledgement: The Scientific Committee wishes to thank the members of the Working Group on 90-day feeding trials on whole food/feed: Gerrit Alink, Gianfranco Brambilla, Noel Albert Dierick, Michael Festing, Gerhard Flachowsky, Gijs Kleter, Joel Guillemain, Harry Kuiper (Chair), Morten Poulsen, Vittorio Silano and Hans Verhagen for the preparatory work on this scientific opinion, and EFSA staff Saghir Bashir, Bernard Bottex, David Carlander, Ellen van Haver, Daniela Maurici and Laura Martino for the support provided to this scientific opinion.

Suggested citation: EFSA Scientific Committee; 2011. EFSA guidance on conducting repeated-dose 90-day oral toxicity study in rodents on whole food/feed. EFSA Journal 2011;9(12):2438 [21 pp.] doi:10.2903/j.efsa.2011.2438. Available online: www.efsa.europa.eu/efsajournal

SUMMARY

Following a request from the European Commission the Scientific Committee was asked to develop principles and guidance for the establishment of protocols for 90-day feeding studies in rodents with whole food/feed. The design of such protocols should be based on the specific properties of food/feed derived from genetically modified plants and other novel food under investigation and in line with the purpose of the study. In view of the multidisciplinary nature of this subject, the task was assigned to the Scientific Committee.

Risk assessment of whole food/feed comprises an integrated approach where information is required on a number of characteristics from various types of tests, including toxicity. Data generated from toxicity testing, whether collected from *in vivo* or *in vitro* studies, provide fundamental information to carry out a risk assessment of a food for human consumption, or of a feed for animals.

This guidance further develops the general procedure set out in the OECD Guideline for the Testing of Chemicals – Repeated Dose 90-day Oral Toxicity Study in Rodents (OECD TG 408), and provides specific advice for performing and reporting experiments carried out with whole food/feed. The guidance aims to help applicants in designing, conducting, analysing, reporting and interpreting repeated-dose 90-day oral toxicity studies of whole food/feed in rodents for the purpose of risk assessment. It is however not intended to provide prescriptive experimental test protocols to carry out such an experiment. It is also outside the scope of the guidance to address the question whether and when such testing is warranted for the safety assessment of whole food/feed.

The conduct of the study and documentation should follow good laboratory practice.

Appropriate characterization of the whole food/feed to be tested is required and should include, among others, a description of the source, its composition, the manufacturing process, information on stability and the presence of chemical and/or microbiological contaminants.

Preparation of appropriate test diets is a key element of the experiment with respect to the choice of the diet type, nutritional balance and necessary adjustments, processing, and storage. Since it is often not possible to include whole foods in an amount that will reliably induce toxicity and thus to obtain a dose-response relationship, fewer dose levels but more animals in control and top dose groups should be used to maximise the power of the study. The use of one control group plus two dose levels (high and low) are recommended when testing whole food/feed. The highest dose level of the whole food/feed should not cause nutritional imbalance or metabolic disturbances in the test animal, and the lowest dose level should always be above the anticipated human/target animal intake level.

In accordance with the European Directive 2010/063, the test animals should be housed socially. The guidance recommends that animals of the same sex are housed in pairs. The experimental unit (ExpU) is therefore a cage containing two animals. More than two individuals per cage are also acceptable if justified, in which case the cage containing more than two animals is the experimental unit.

A comprehensive set of endpoints as set out in the OECD TG 408 should be measured during and at the end of the 90-day period, as appropriate; additional endpoints may be considered for the assessment, depending on the nature of the food/feed being tested and the available information. A comprehensive statistical analysis should be conducted including the analysis of the differences between males and females.

A randomised block design is recommended for 90-day toxicity studies when testing whole food/feed. Other designs may be acceptable provided that appropriate justification for using them is given.

The use of a power analysis to estimate a sample size capable of detecting a pre-specified biologically relevant effect size with a specified power and significance level should be done to determine an

appropriate sample size. Those objectives have to be specified clearly and translated into effect sizes that are considered by experts as being biologically relevant. The biological relevance of any observed differences should be considered, whether or not they reach the chosen level of statistical significance. This assessment should involve the use of point and interval (e.g. confidence) estimates in addition to the significance level.

In the case of GM food, the inclusion in the experimental design of reference groups, fed with a diet containing commercially available food/feed similar to the test food/feed, in order to estimate the natural variability of endpoints is in general not recommended since this would substantially increase the number of test animals. Historical background data on variations in endpoint values should primarily be obtained from databases available in the actual testing facility or in the public domain. Inclusion of reference groups should be considered if no acceptable historical background data is available.

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY THE EUROPEAN COMMISSION

Evaluation of the safety and nutritional properties of whole genetically modified (GM) and other novel foods/feeds is an important feature in the safety/nutritional assessment of these foods/feeds (Regulation (EC) No 1829/2003 on GM food/feed and feed and Regulation (EC) No 258/97 on Novel Foods under revision).

Commonly the safety assessment of these foods/feed comprises an extensive compositional analysis, an *in vitro*/*in silico* characterisation and assessment of results obtained from animal tests with relevant purified compounds identified in them, like for instance newly expressed proteins or other constituents, rather than the toxicological/nutritional testing of the whole food/feeds themselves. In specific cases toxicity testing of the whole food/feed may be considered, depending on the type of the food/feed under investigation, its history of (safe) use, the available toxicological information, or remaining uncertainties. As of today, no standardised protocol or guidelines exist for this type of study and applicants are advised to adapt the OECD Test Guideline 408 (90-day oral toxicity study in rodents) designed for toxicity testing of single defined substances.

In March 2008, a report of the EFSA GMO Panel Working Group on animal feeding trials entitled "Safety and nutritional assessment of GM plants and derived food and feed: The role of animal feeding trials" was published. This publication treats this issue in more detail and recommends the development of supplementary guidelines for this type of study.

It is appropriate that EFSA develops further guidance for applicants on this matter. This work could also contribute to further harmonising the current performance of animal feeding trials with whole food/feed at the international level.

TERMS OF REFERENCE AS PROVIDED BY THE EUROPEAN COMMISSION

EFSA is requested according to Article 29 of Regulation (EC) No 178/2002 to develop principles and guidance for the establishment of protocols for 90-day feeding studies in rodents with whole food/feed. The design of such protocols should be based on the specific properties of the GM and other novel food/feed under investigation and in line with the purpose of the study. Specific attention will be paid to the development of protocols suitable for food/feed derived from GM plants.

Guidance should include among others considerations on:

- Study purpose and design,
- Type of test, control and reference diets, analysis and storage,
- Criteria for balancing diets,
- Types of test, control and reference groups,
- Dosage regimes and spiking,
- Toxicological and nutritional endpoints to be measured,
- Data collection, statistical analysis,
- Quality assurance aspects.

ASSESSMENT

1. INTRODUCTION

Risk assessment of whole food for human consumption or of whole feed for animal use is currently based on integrated approaches where information, generated through various types of tests, is required on a number of characteristics and properties including source material, production and processing methods, chemical composition including contaminants as well as nutritional and toxicological properties. In such a framework, relevant toxicity data may be obtained from *in vivo*, *in vitro* and/or *in silico* studies.

The OECD Guideline for the Testing of Chemicals – Repeated Dose 90-day Oral Toxicity Study in Rodents (OECD TG 408) aims at providing an understanding of possible adverse effects due to repeated exposure of test animals to chemicals over a prolonged period of time (90-days) covering post-weaning maturation and growth into adulthood (OECD, 1998a). The OECD TG 408 is designed to provide treatment-related information on toxic effects observed in test animals, with identification of target organs, and possibly mode of action as well, to identify a no-observed-adverse-effect level. Compared with the original guideline adopted by OECD in 1981, the current version of the OECD TG 408 places additional emphasis on neurological endpoints and provides indications also on immunological and reproductive effects.

The main difference between testing chemicals and whole food/feed is that chemicals can be administered to the test animals at dose levels which are much higher than the likely human exposure levels, whereas such a testing approach is not always possible with whole food or feed as they are bulky. In fact, administering high dose levels of whole food/feed is likely to result in satiation and/or unbalanced diets. Therefore, careful consideration should be given to effective ways in which the design, conduct and analysis of the OECD TG 408 are adapted to specific whole food/feed testing in order to increase the chance of detecting any toxicologically-relevant effects. However, in some cases, the administered dose level can be equivalent to a high level of human consumption.

During the development of the current guidance, attention has been paid to the recommendations for carrying out statistical analyses of data from 90-day rat feeding studies in the context of marketing authorisation applications for genetically modified (GM) organisms, recently published as a contribution to EFSA by the French Agency for Food, Environmental and Occupational Health and Safety (ANSES, 2011)

A draft of the guidance underwent a public consultation from the 7th of July until the 22nd of August 2011. The comments received were considered and have been incorporated where appropriate. A technical report of EFSA on the outcome of the public consultation including the comments is available on the EFSA website.

1.1. Objective of the guidance

This guidance aims to help applicants in designing, conducting, analysing, interpreting and reporting repeated-dose 90-day oral toxicity studies of whole food/feed in rodents (generally rats and mice) for the purpose of risk assessment. The guidance draws attention to possible approaches to: (i) minimise bias, (ii) maximise the precision and power of the experiment, (iii) strengthen the statistical analysis of experimental data, and (iv) improve reporting and analysis of the results. It is not the objective of this guidance to provide prescriptive experimental test protocols to carry out such a test because methods of testing may depend on the properties of individual feed/foods. Testing of the safety of the whole food/feed should be considered based on hazards identified during the preceding molecular, compositional, phenotypic or agronomic analysis. The objective of a 90-day feeding trial is to detect any possible toxicological effects of the test diet compared with the control diet.

It is outside the scope and aim of this guidance to address the question whether and when such testing is warranted for the safety assessment of whole food/feed. For GM food and feed, this subject has already been addressed elsewhere (EFSA GMO Panel Working Group on Animal Feeding Trials, 2008; EFSA Panel on Genetically Modified Organisms (GMO), 2011a).

For novel foods, it is well established that the safety assessment of a large number of novel foods for which a traditional counterpart exists, has been based on the acceptance of substantial equivalence of the novel food with already existing foods in terms of their composition, nutritional value, metabolism, use modalities, nature and levels of undesirable substances (e.g. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1997R0258:20090120:EN:PDF and http://ec.europa.eu/food/food/biotechnology/novelfood/notif_list_en.pdf#page=71).

1.2. The term "whole food/feed" as used in the guidance

In the context of this guidance, "whole food/feed" refers to a product to be consumed by humans or a feed material consumed by animals, which is composed of a multitude (up to thousands) of individual substances. Whole food/feed range from plant based products such as maize or potatoes to more refined products such as fruit juices or flour, to foods and feed consisting of microorganisms as well as animal-derived food products such as meat and milk.

The interpretation of the whole food/feed term as used in this guidance (sometimes also referred to as "whole product") aims to differentiate a whole food/feed from more purified single food/feed ingredients or fractions that, in the context of animal testing, could be administered at much higher dietary levels. As such, individual ingredients are outside the scope of this guidance.

1.2.1. Food/feed derived from genetically modified organisms

Products under consideration include whole food/feed derived from GM plants with various alterations to introduce e.g. herbicide tolerance and/or insect resistance (including stacking of such events) and traits leading to improved responses to environmental stress conditions, or to improved nutritional/health characteristics (see further Table 1 of the EFSA Report on Animal Feeding Trials, 2008). Typical GM crops are maize, soybeans, oilseed rape and cotton. This category also includes genetically-modified microorganisms and their products.

Furthermore, products under consideration may be derived from GM animals whose genetic material has been altered in a heritable way either through recombinant DNA or other *in vitro* nuclear techniques. Applications may include genetic modification of husbandry animals, fish, as well as crustaceans and molluscs, insects (for instance honey bees) and other invertebrates. Inserted traits can be related to more efficient or increased production of food, enhanced nutritional characteristics and wholesomeness of these foods, lower emissions to the environment or improvement of the health characteristics of the GM animal, including better resistance to abiotic stressors and pathogens, improved fertility and lower mortality.

1.2.2. Novel foods

Products under consideration are novel foods falling under Regulation (EC) No 258/1997. This may be the case for: (i) whole novel foods consisting of microorganisms, fungi or algae; (ii) whole novel foods consisting of parts of new plants or animals; and (iii) whole novel foods to which a production process not currently used has been applied, where that process gives rise to significant changes in the composition or structure of the foods which affect their nutritional value, metabolism or level of undesirable substances. The full list of currently authorised novel foods in the EU is found at: http://ec.europa.eu/food/food/biotechnology/novelfood/authorisations_en.htm. Whereas individual ingredients are out of the scope of this guidance, it is noted that the majority of novel food products consists of novel food ingredients rather than whole novel foods.

2. CHARACTERIZATION OF THE WHOLE FOOD/FEED AND PREPARATION OF THE DIET

2.1. Characterisation of the whole food/feed

Before preparing the animal diet, it is necessary to have a complete analytical picture of the composition of the whole food/feed and the comparator (in the case of GM foods). The whole food/feed analysed should be that which is to be incorporated in the diet of the test animal, and relevant to the product to be consumed. The compositional analyses should include macro- and micronutrients, other food/feed constituents, and chemical and microbiological contaminants.

Critical elements for the characterisation of food/feed have been described in various documents (Regulation (EC) No 258/97; Verhagen et al., 2003; Aggett et al., 2005; EFSA Panel on Genetically Modified Organisms (GMO), 2011a). The following are examples of elements that should be addressed to obtain a complete analytical composition of the whole food/feed: name, source and specifications, composition, manufacturing processes, batch to batch variation, information on stability etc (for specific details, see EC Scientific Committee on Food, 2001). For GMOs, this would also include molecular characterisation of the genetic modification.

2.2. Choice and preparation of the diet

The performance of laboratory animal studies on whole food/feed products as covered in this guidance meets a number of challenges since they are complex mixtures of compounds with very different biological characteristics. Food/feed are bulky; they may have an effect on the satiety of animals and can therefore only be fed at relatively low multiples compared to their typical presence in the human/target animal diet. Moreover, there is a possibility that in attempting to maximise the dietary content of the whole food/feed under investigation, nutritional imbalances may occur. These could lead to the appearance of effects which may not be related to the properties of the whole food/feed being tested.

2.2.1. Choice of diet

There are several types of diets in which the whole food/feed to be tested could be incorporated with to form the animal test diet. The most common diets in animal studies are the following:

- Diets based on natural ingredients,
- Purified diets (formerly known as semi-synthetic diets)
- Synthetic diets which are chemically designed
- Human-type diets

Natural-ingredient diets are formulated with cereals, legumes and other food/feed products. They are acceptable and palatable to most animal species. They include the commercially produced standard laboratory animal diets, known as chow diets, which are often used for rodent feeding studies testing chemicals. A natural-ingredient diet is often used for the testing of whole GM food.

Purified diets are formulated with a more refined and restricted number of ingredients than the natural-ingredient diet. The ingredients are well-characterised and may include maize starch, soy oil, sucrose, casein, cellulose etc. Purified diets can be used when foods and macro-ingredients are tested because it is easy to alter ingredients in this type of diet and it is possible to achieve a higher incorporation level of the whole product to be tested than in the natural-ingredient diet.

Synthetic diets are made from simple ingredients like amino acids and specific fatty acids and are used to test single chemically defined substances like a specific micronutrient or amino acid. The synthetic diet is expensive and rarely used.

A human-type diet should represent a balanced human meal and at the same time fulfil the nutritional requirements of the experimental animal. This type of diet is not used very often due to the lack of background experience and its complex nature.

2.2.2. Formulation of nutritionally balanced test diets

A high incorporation level of a whole food/feed in the diet of the test animals can result in a nutritional imbalance. These effects can be predicted from the compositional analysis, review of literature or preliminary studies and should be taken into account in the test diet formulation. The anticipated level of nutritionally important ingredients in the diets of the control and dose groups should be examined. In order to avoid nutritional imbalance or metabolic disturbance (NRC, 1995), adjustments should be made for differences in composition greater than 5% (FDA, 2000; Knudsen and Poulsen, 2007; Poulsen et al., 2007).

When testing complex novel products of protein origin or with a high content of protein it is frequently the protein *per se* that is the limiting factor in the attempt to get as high an incorporation level as possible. If a significant amount of the protein in the diets is substituted, amino acid imbalances and potential consequences on animal health should be taken into account. For novel fats or products containing high levels of fats it is correspondingly the fat that is the limiting factor that in excess can cause an unbalanced diet. When testing meat-based products, consideration should be given to the fact that rodents, albeit omnivores, are not adapted to a full meat-based diet.

Before preparing a nutritionally balanced diet, consideration should be given as to whether or not the whole food/feed contains inherent anti-nutritional components in a relatively high concentration (e.g. trypsin inhibitor in unprocessed soybean meal or glyco-alkaloids in potatoes). These effects can be predicted from the compositional analysis, review of literature or preliminary studies and should be taken into account in the test diet formulation. The presence of anti-nutritional components, or other substances, in the whole food/feed to be tested may be the limiting factor that determines its maximum inclusion level into the test diet, or otherwise influences the results of the test.

2.2.3. Choice of control diets

Guidance for the selection of a suitable control for GMOs has been described in the EC Regulation No 1829/2003 on genetically modified food and feed[4] which defines a suitable control (conventional counterpart) as "*similar food or feed produced without the help of genetic modification and for which there is a well-established history of safe use*". The EFSA GMO Panel has required the use of non-GM lines with comparable genetic background, i.e. near-isogenic in the case of sexually propagated crops or isogenic in the case of vegetatively propagated crops (EFSA Panel on GMO, 2011b), as comparators in its evaluation of GMO applications. Control diets for novel foods should be selected on a case-by-case basis, depending on the nature of the novel food to be tested.

2.2.4. Influence of processing of the whole food/feed in the test diets

The safety profile of the whole food/feed may be changed by processing (e.g. in case of presence of thermolabile toxic substances). This should be taken into account when preparing the test diet.

2.2.5. Analysis and storage of the test diets.

The compositional analyses of the test diets should include macro- and micronutrients, and other food/feed constituents.

The potential occurrence of biological and chemical contaminants in the test diet should be reported. Acceptable levels in rodent diets have been issued by different national bodies, framed within Good Laboratory Practices (GLP) guidelines (Clarke et al., 1977; Rao and Knapka, 1987; Stevens and Russel, 2007; Directive 2010/63/EU).

[4] See http://ec.europa.eu/food/food/animalnutrition/labelling/Reg_1829_2003_en.pdf

Good manufacturing techniques and appropriate environmental storage conditions will minimize spoilage and degradation of the test diets. Guidance on how to store feeding stuffs, preventing nutrient degradation and mould and insect colonisation and growth, are implemented under Hazard Analysis Critical Control Points schemes (e.g. TASCC, 2010). The homogeneity and stability of the test diets under the conditions of administration should be determined and reported.

3. RODENTS USED IN 90-DAY TOXICITY STUDIES

The general principles for using laboratory animals should be adhered to. All studies should be carried out following OECD Good Laboratory Practice (GLP) guidelines (OECD, 1998b) and taking account of animal welfare as outlined by the EFSA Scientific Panel on Animal Health and Welfare (AHAW) opinion related to the aspects of the biology and welfare of animals used for experimental and other scientific purposes (EFSA, 2005) and the EFSA Scientific Committee opinion on existing approaches incorporating replacement, reduction and refinement of animal testing: applicability in food and feed risk assessment (EFSA, 2009). All procedures should conform with Directive 2010/63/EU of the European Parliament and of the Council or similar legislation in non-EU countries, and as such take account of the "3Rs" (Replacement, Refinement and Reduction) .

The studies should use both sexes of either rats or mice. Rats have a long history of use in toxicity testing and are widely preferred by toxicologists. They are easy to handle and their large size ensures that ample tissues and body fluids are available. However, the use of the mouse should not be precluded. They are widely used in absorption, distribution, metabolism and excretion (ADME) studies; Animals used in 90-day toxicity experiments should be healthy and free of the major pathogens (i.e. of specific pathogen free (SPF) status according to FELASA guidelines (FELASA, 1996). They should come from breeding colonies maintained to internationally recognised standards such as AAALAC (http://www.aaalac.org/) accreditation or its equivalent, with a routine health monitoring system which screens for pathogenic bacteria, viruses and parasites. The list of pathogens tested in the screening should be included in the study report, stating those which were present and absent in the breeding colony from which the animals were obtained.

At the commencement of the study, the weight variation of animals should be minimal and not exceed ± 20 % of the mean weight of each sex, when using a completely randomised design (OECD TG 408). When a randomised block design is used, test animals within a block should be matched for body weight as closely as is practically possible (e.g. ± 5 % of mean body weight of the block).

3.1. Choice of stocks or strains of animals

There are two major classes of laboratory mice and rats used in research and testing: outbred stocks and isogenic strains (inbred and F1 hybrid). Outbred "genetically undefined" stocks such as "Sprague-Dawley" are produced in closed colonies where each individual is genetically unique, and rats from different breeders will be genetically different. Animals from such stocks tend to be phenotypically more variable than isogenic strains and the colony is less stable. In contrast, inbred strains and F1 hybrids (the first generation cross between two inbred strains) are "genetically defined". They are stable over time and cannot be changed by selective breeding, although there are sub-lines of many of the most widely used inbred strains. Further details of these two classes of stock and strains and the genetic nomenclature rules are given elsewhere (Festing and Lutz, 2010, 2011).

Most toxicity testing of foods, food constituents, food additives or food contaminants is carried out using outbred strains. Chosen stocks or strains should be designated according to internationally accepted nomenclature rules[5]. Studies based on either inbred or outbred stock can be acceptable but the reason for choosing a particular strain or stock should be stated.

[5] See http://www.informatics.jax.org/

3.2. Housing and maintenance

Rats and mice are social animals and there is substantial literature showing that housing them singly causes stress (Lapiz et al., 2001; Myhrer, 1998; Westenbroek et al., 2003; Leshem and Sherman, 2006). Point 3.3 of Annex III of the European Directive 2010/63/E states that "animals, except those which are naturally solitary, shall be socially housed in stable groups of compatible individuals … unless reasons for derogating are provided".

It should be noted that the number of animals per cage has statistical implications (see section 6.1.4).

4. ENDPOINTS TO BE MEASURED

The 90-day study in rodents should be conducted with the full range of observations as described in the OECD TG 408. Additional endpoints described in the more recent guideline on repeated-dose 28-day oral toxicity study in rodents (OECD TG 407) may be considered for assessment, depending on the nature of the food/feed being tested and the available information.

Other endpoints should also be considered if there are indications that the whole food/feed may have effects on e.g. the cardiovascular, nervous, gastrointestinal tract or immune system. For instance, if the whole food/feed is expected to have an impact on the gut, then the microbial flora should be investigated. Additional markers of potentially adverse nutritional and/or metabolic effects should be considered on a case-by-case basis, according to the available body of evidence and the type of whole food/feed under investigation.

The endpoints should be examined for all animals, except for histopathology which initially should be performed on the control and high dose group (in case of novel foods), or on the high dose control and the high dose test group (in case of GM food/feed) (see section 6.1.3). If histopathological differences between test and control groups are observed, those from other groups should also be examined.

5. STUDY CONDUCT AND DOCUMENTATION

The study conduct and documentation should follow good laboratory practice as described by OECD (OECD, 1998b).

6. EXPERIMENTAL DESIGN AND STATISTICAL METHODS

There is extensive literature on methods of designing and analysing randomized experiments (Fisher, 1960; Cox, 1958; Cochran and Cox, 1957; Montgomery, 1984; Mead, 1988). In this section experimental design and statistical methods are addressed. The intention is to encourage the use of efficient designs and appropriate statistical methods. The methods below offer flexibility and are not intended to be prescriptive. The key objective is to obtain accurate and unbiased estimates of the food/feed effect. If this objective is met then these estimates can be compared with historical data and any future results, provided that they are also unbiased.

Experiments and the statistical analysis should always be pre-planned. Additional groups should not be added during the course of the experiment.

6.1. Experimental design considerations

6.1.1. Objectives

The objectives of the experiment should be clearly stated. It should be designed to be unbiased, with no systematic differences among groups apart from the treatment. This is mainly controlled by appropriate randomisation and by using coding (where practical) to blind the staff to the treatment group to which an experimental unit (see section 6.1.4) belongs. It is noted that, in many cases, pathological assessment will initially be done un-blinded followed by a blind re-assessment, should there be evidence of a treatment effect (Crissman et al. 2004).

The experiment should be powerful: if there is a true difference between the treatment groups, then the experiment should have a good chance of detecting it, although false negatives (beta errors) and false positives (alpha errors) can never be eliminated. Power depends on controlling inter-individual variation, on the magnitude of the difference between the treated groups, on sample size, and on the acceptable levels of false positive (usually set at 5 %) and false negative (often set at 10-20 %) results. The experiment should also have a wide range of applicability. For example, sex-dependent effects may be present, so both sexes should be included.

6.1.2. Completely randomised design versus randomised block design

In a completely randomised design, the experimental units are assigned to treatment groups at random regardless of their characteristics. Completely randomized designs are simple and can easily accommodate unequal numbers in each group. Their disadvantage is that if the experiment is relatively large they may become difficult to handle without introducing unwanted sources of variability. For example, it may be difficult to obtain a large number of test animals of uniform weight and age, house them all under identical conditions in different cages and gather data from them all over a short period of time. The introduction of sources of variability such as the above-mentioned ones can produce confounding and reduce the power of the experiment. To overcome some of these issues a randomised block design can be used.

The randomised block design involves splitting the experiment up into a number of "mini-experiments" or blocks, which are then re-combined in the statistical analysis. Animals are matched within blocks according to factors that generate variability without being of direct interest (e.g. body weight, location of the cages, timing of making the measurements/determinations of the end points). The aim of this matching is to control for the unwanted sources of variability. Each block usually contains one experimental unit of each of the treatment groups; for example, if there are 8 treatment groups, then each block will consist of 8 experimental units, one on each of the treatments. In the majority of cases randomised block designs are more powerful than a completely randomised design (Mead, 1988). They can reduce the number of experimental units needed whilst keeping the same power and confidence level, or increase the power for a fixed sample size. For this reason they have been widely used especially in agricultural and industrial research (Mead, 1988). A randomised block design is recommended instead of a completely randomised design for 90-day food/feed studies when testing GM and novel foods. Other designs may be acceptable provided that appropriate justification for using them is given.

6.1.3. Dose, Reference groups and historical control data

Two dose levels (in addition to the controls) high and low are recommended when testing whole food/feed. The high dose level should correspond to the highest level of the whole food/feed that can be incorporated in the animal diets whilst avoiding nutritional imbalances. The low dose level could be half to a quarter of the high dose and should always be above the anticipated human intake, as otherwise the data obtained will be of no relevance in the assessment. The applied dose levels may be impacted by processing and this should be taken into account.

The whole food/feed to be tested should preferably be fed *ad libitum*. In the case when the whole food/feed is given alongside the diet, or by gavage, the same kind of consideration about balancing the diet should be taken as in the case when the whole food is incorporated into the diet. Administration by gavage is not common for whole food/feed but could be considered in certain instances due to poor palatability or stability, or in cases where an exact dosing is needed.

For whole food/feed where no adequate information exists, a small preliminary tolerance test with a limited number of animals and with a short duration (1-2 weeks) could be conducted on a case by case basis. Pilot studies may be conducted on a case-by-case basis to investigate the appropriateness of the dose levels to be used in the 90-day study.

When testing GM foods, isogenic (or near isogenic) controls should be used for comparison with the GM food to avoid confounding due to the different composition of the diet. For instance, if diets with 15% and 30% GM maize are administered to treatment groups then diets with the same percentages of non-GM maize of the same strain (the isogenic control) should be used. Another approach could consist of three test groups, for example, an isogenic control group with, say, 30% maize, a group with 15% isogenic control plus 15% GM maize, and a group with 30% GM maize. For novel foods, in many cases, two dose levels in addition to the control would be appropriate.

In the case of GM food, the inclusion in the experimental design of reference groups, fed with a diet containing commercially available food/feed similar to the test food/feed, in order to estimate the natural variability of endpoints is in general not recommended since this would substantially increase the number of test animals. Historical background data on variations in endpoint values should primarily be obtained from databases available in the actual testing facility or in the public domain. Inclusion of reference groups should be considered if no acceptable historical background data is available.

6.1.4. Specification of the experimental unit as a cage

This guidance recommends that animals of the same sex should be housed in pairs. In a controlled experiment, experimental units (ExpU) are assigned to the treatments at random, and it must be possible for any two ExpU to receive different treatments. Animals in the same cage cannot receive different treatments when these are supplied in the diet. However, cages can be independently assigned to treatments, so a group of animals within a cage is the ExpU.

The power of the experiment is largely determined by the number of ExpU. Housing more than two animals per cage will increase the total number of animals with little scientific benefit because it will not increase the number of cages. Housing the animals one per cage or more than two per cage should only be done if there is a convincing scientific case for doing so.

If there is more than one animal assigned to a cage receiving the same treatment, the statistical analysis cannot be done assuming that the animal is the ExpU (i.e. done on individual animals), as the significance levels for the statistical testing may be incorrect.

The statistical analysis should first test whether there are statistically significant (usually at $p<0.05$) cage effects, i.e. differences in the individual endpoints such as blood cell counts between cages having the same treatment. If not (i.e. the between-cage variation is not statistically significantly greater than the between-animal variation), then the statistical analysis can be based on the individual animals. However, if there are significant cage effects, then the analysis should be based on the mean of the two animals within each cage.

Animals should be individually identified and all data should be collected separately for each animal, except for food and water consumption.

6.1.5. Determination of sample size and power

The estimates of the power of the experiments presented in this section assume that the animals are housed two per cage.

The required sample size in clinical trials is usually determined using a power analysis. This requires a decision on an effect size of scientific interest (EFSA, 2011) and an estimate of the standard deviation (assuming quantitative variables). However, there are difficulties in defining the effect size in a toxicity test where there are multiple possible outcomes, any one of which could indicate toxicity. Therefore, sample size in toxicity testing protocols such as OECD 408 has largely been based on past experience.

An alternative approach described here, is to base sample size on a pre-specified effect size measured in SD units. This is known as the "standardised effect size" (SES). It is the difference between treatment groups divided by the standard deviation (SD) among experimental units, and can be regarded as a signal/noise ratio. Power analysis can then be used to estimate the required sample size needed to be able to detect a specified effect size in these units. If experience from previous toxicity tests shows that an effect size of, say, one SD or less is of little toxicological relevance then this can be used to determine sample size in new situations.

Using this method, an appropriate sample size (number of ExpUs) can be estimated from the SES, the significance level (usually set at 5 %), the power (often set at 80-90 %) and the alternative hypothesis (one or two sided). The relationship among these variables is shown graphically in Figure 1, where the effect size is specified in terms of the SES and a two-sided alternative hypothesis assumed. The assumptions used for the calculations of the graphs in Figure 1 do not include a sex by treatment interaction. However, the analysis should investigate sex by treatment interactions, or any other interactions that are deemed to be of importance. It does not take account of blocking, which may increase the power by an unknown amount.

As examples, experiments designed to detect an SES of one SD or less, whilst aiming to achieve 80-90 % power are presented in Table 5.1 and further discussed below.

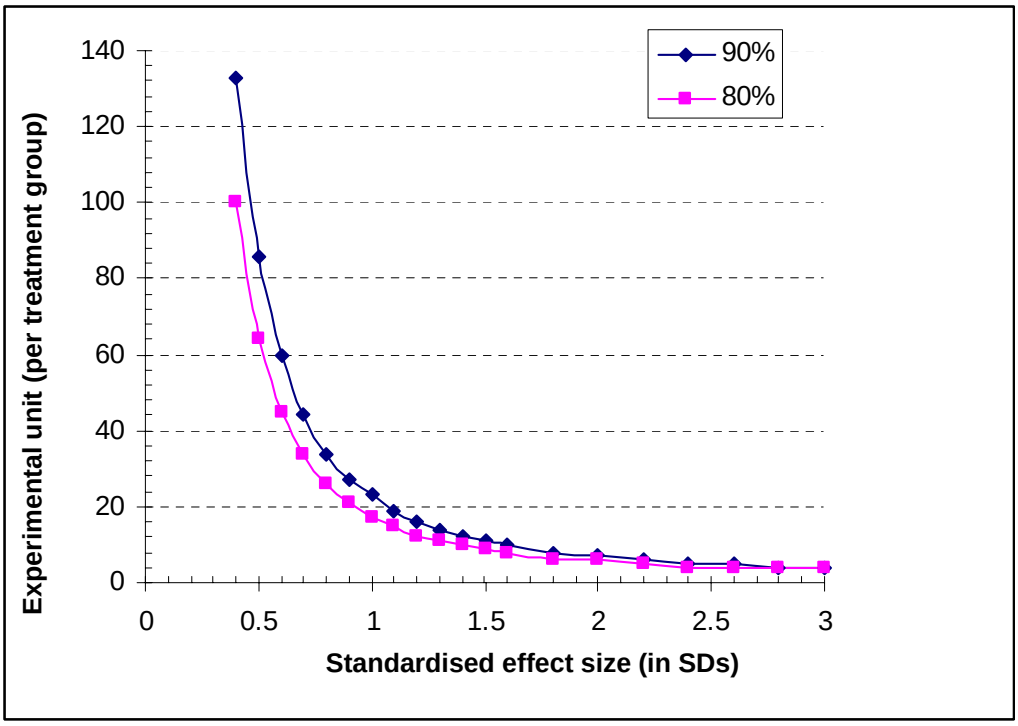

Figure 1: Number of experimental units needed per treatment group as a function of Standardised Effect Size for an 80 % and 90 % power and 5 % significance level using a two-sided t-test. This will approximate the situation in a 2 (treatments) x 2 (sexes) factorial design.

For the testing of chemicals, OECD 408 specifies the use of 80 animals comprising both sexes in four treatment groups (control, low, medium and high dose). As animals can be individually dosed, each animal is an experimental unit. It does not specify whether a completely randomised or a randomised block design should be used, but a completely randomised design is assumed. Row A of Table 5.1 (and Figure 1) shows that it has an 80 % chance of detecting a standardised effect size of 0.9 SD and a

90 % chance of detecting an SES of 1.1 SDs assuming no sex by treatment interaction. These calculations are based on the number of experimental units in the control and top dose group.

As stated earlier, a cage with two animals is the experimental unit. One effect is that the SD among experimental units will be reduced by a factor of the square root of two, so the number of cages can be reduced (i.e. putting two animals in each cage does not mean that twice as many animals are needed). The number of treatment groups could be reduced compared to the recommendations in OECD TG 408, so that the group size in the controls and top dose groups can be increased, without increasing the total number of animals. Further increases in power may be obtained by using a randomised block design, although this cannot be quantified as it depends on the heterogeneity between the blocks.

Table 5.1 gives four examples of experimental plans using various numbers of animals and treatment groups which all have sufficient power to detect an SES of approximately one SD or less. Example A relates to the OECD TG 408 completely randomized design with a cage with one animal being the experimental unit; this can detect an effect size of about one standard deviation. Examples B to D make an adjustment for having two animals per cage. Example B shows a plan with 96 animals, two animals per cage, and a randomised block design where there are three treatment groups (control, low and high). Example C shows a plan with four treatment groups (e.g. 15% GM maize versus 15% isogenic control and 30% GM maize versus 30% isogenic control) in which the statistical analysis is done as a 2^3 factorial design, i.e. 2 genotypes (GM and control) X 2 sexes X 2 doses. Example D shows a plan similar to Example B which is more powerful, i.e. it can detect a smaller effect size by using 144 animals.

Table 1: Examples of detectable effect sizes with different designs and numbers of experimental units assuming a two-sided t-test involving the control and top dose group, a 5% significance level, power as shown with the two sexes pooled

Experimental plan	Design[(a)]	Total number of animals	Number of animals per cage (n)	Number of treatment groups	Number of blocks	Number of ExpU per block	Number of ExpU per treatment group	Estimated effect sizes (SDs)	
								80% power	90% power
Example A (OECD 408)	CR	80	1	4			20	0.9	1.1
Example B	RB	96	2	3[(c)]	8	6	16 (23)[(b)]	0.9	1.0
Example C	RB	96	2	4[(d)]	6	8	24 (34)[(b)]	0.7	0.8
Example D	RB	144	2	3[(c)]	12	6	24 (34)[(b)]	0.7	0.8

[(a)] CR is completely randomised, RB is randomised block
[(b)] Number of singly housed equivalents (number of experimental units times the square root of "n"), on which power calculations are based
[(c)] For testing novel or GM foods with three treatment groups (control, low, high)
[(d)] For testing GM foods with four groups (see text).

The power of the experiment can be increased by using a higher significance level than 5 % which is the statistical level most commonly used in biological research. However, a higher statistical significance level will also increase the number of false positive results. During the assessment, emphasis should be placed on the biological relevance of any observed differences whether or not they reach the chosen level of statistical significance (EFSA, 2011). This is best done by looking at the point and interval (e.g. confidence) estimates and not by just focussing on the P-value.

All the data should be analysed together, i.e. for both sexes, to maximise the power. Treatment effects for each sex should be estimated separately only in the presence of treatment by sex interactions.

In the protocol the applicant should justify the sample size calculation including the variables (i.e. SES, significance level, power and the alternative hypothesis). In addition, the design of the experiment should also be clearly described including whether it is a "completely randomised design" or a "randomised block design" and the ExpU should be specified (e.g. number of animals/cage).

6.2. Specification of the methods of statistical analysis and reporting

All the important details about the experimental design and an overview of statistical methods, including the design, and analysis, should be documented in a protocol prior to the start of the trial. The statistical analysis plan (SAP) should be written and signed off prior to the end of the experiment. A statistical report should be written with all the analysis results as documented in the SAP. Any supplementary or additional analysis should be justified and documented. The statistical analysis programs, logs and outputs should be provided for the purposes of the review. The data should be made available on request in a suitable electronic format with table and column definitions.

Steps in the statistical analysis should include the screening of the data for outliers, transformation of scale where necessary and the choice of the most appropriate statistical tests taking account of the distribution of the observations.

Descriptive statistics should be presented for all environmental and analysis variables (endpoints). The summary statistics should include the mean, standard deviation, median, minimum and maximum. The use of graphical methods such as plots of means and 95 % confidence intervals for each group, and/or box and whisker plots is encouraged. All individual raw data should be provided in a suitable electronic format (e.g. XML, csv, etc.) for eventual further analysis. Table and column definition should be also provided.

The results from the statistical analysis should be presented in the original units and in terms of the standardised effect size using point and interval estimates (e.g. confidence). The standardised effect sizes (differences between treatment groups)/SD with 95 % confidence interval (Nakagawa and Cuthill, 2007; Ellis, 2010) for all endpoints can be shown on the same graph. The biological relevance of all statistically significant differences (EFSA, 2011) as well as the point and interval (e.g. confidence) estimates of any responses (some of which may not reach statistical significance) should be considered by a qualified toxicologist (see section 6).

The statistical analysis plan should detail all the analysis methods with all the results reported in the final report. Sex-specific traits (i.e. ones such as testis and uterus weights) which can only be measured in one sex should be analysed using an appropriately reduced analysis of variance. Any statistically significant interactions, particularly those involving treatment and sexe should be fully explored. Where the statistical analysis is conducted by sex, the results should be presented consistently for each sex and for both sexes combined to assist the risk assessor.

The separate analysis of other endpoints, most of which are not expected to differ between treatment groups, may result in a large number of statistical tests. This will lead to the issue of multiple testing (multiplicity) and therefore it should be addressed by the applicant in the protocol, statistical analysis plan and study report. Any methods used to adjust for multiplicity should also be clearly documented and referenced. With a randomised block design, block is also a random factor which should be included in the model.

The protocol should describe the intended methods of statistical analysis and the methods employed to minimise the bias. The statistical analysis should provide the full details of the intended analysis including full descriptions of the statistical models fitted.

The reporting of the statistical analysis should be consistent with the protocol and the statistical analysis plan. Any unplanned analysis should be justified and detailed in the final study report. The results should be presented in a consistent and clear manner to facilitate the interpretation by the risk assessors.

7. INTERPRETATION OF RESULTS OF ANIMAL STUDIES

Any effects observed in the animals should be evaluated in the context of the whole animal's physiology in order to assess their relevance for the safety of the whole food for humans or of the whole feed for target animal species.

7.1. Possible dose-related trends

When differences in endpoints are observed between animals on the test and control diets, these should be evaluated with respect to type and magnitude of the findings, frequency, possible trends, natural ranges of variability, and possible correlation with other changes, in order to determine whether a treatment relation may exist or whether the observed differences are accidental or spurious. Specific attention should be paid to situations where effects are only observed at low dose levels. Supportive data for possible causality between the test food/feed and effects in test animals may include additional toxicity data obtained in other studies.

7.2. Possible interrelationships between endpoints

A change in an isolated endpoint (for instance haematology, biochemistry and pathology) is often of limited interest and the conclusion on biological relevance depends on several endpoints. Observed changes in individual endpoints may be correlated, thus strengthening the indication that an effect has occurred as a result of the treatment. For example, liver damage, observed as a change in histopathology, gross pathology, or organ weights, may also be evident from changed levels in serum of liver-derived enzymes, or bilirubin. Detection of changes in blood parameters by hematological analysis may be interlinked with results from the analysis of bone marrow, spleen, lymph nodes and mononuclear phagocyte system (reticuloendothelial tissue) of various organs and tissues.

7.3. Occurrence of effects in both sexes, reproducibility and animal species specificity

Effects often occur in both male and female animals, but in certain cases one sexe may be more sensitive due to differences for example in detoxification mechanisms, or due to hormonal differences. Differences observed in treated animals may also have been observed in other studies in the same or in another animal species, and it should be realized that certain effects may be specific for the animal test species but not of relevance for humans (idiosyncrasy).

7.4. Background range of variability

Changes in endpoints of treated and non-treated animals must also be evaluated with respect to normal biological variation of the endpoints. As described above, information on the background variability may be obtained from historical control data and, in case such data are not available, from reference control groups.

8. ASSUMPTIONS AND UNCERTAINTY ANALYSIS

A specific chapter on assumptions and uncertainty analysis should be included in the study report. Any uncertainties (in addition to natural variation in biological endpoints) in the design of the experimental model which might influence the power of the experiment should be highlighted and quantified as far as possible. The assumptions underlying the statistical analysis should be reported and tested for robustness.

REFERENCES

Aggett, P.J., Antoine, J.M., Asp, N.G., Bellisle, F., Contor, L., Cummings, J.H., Howlett, J., Müller, D.J.G., Persin, C., Pijls, L.T.J., Rechkemmer, G., Tuijtelaars, S., Verhagen, H., 2005. Passclaim. Eur. J. Nutr. 44, 5–30.

ANSES (Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail), 2011. Opinion of the French Agency for Food, Environmental and Occupational Health and Safety - Recommendations for carrying our statistical analyses of data from 90-day rat feeding studies in the context of marketing authorisation applications for GM organisms. Request no. 2009-SA-0285.

Clarke, HE, Coates ME, EVA JK, Ford DJ, Milner CK, O'Donoghue PN, Scott PP and Ward RJ, 1977. Dietary standards for laboratory animals: report of the Laboratory Animals Centre Diets Advisory Committee. Laboratory Animals 11, 1-28.

Cochran WG and Cox GM, 1957. Experimental designs. New York, London: John Wiley & Sons, Inc., 1-611.

Cox DR, 1958. Planning experiments. Ed. John Wiley and Sons: New York.

Crissman JW, Goodman DG, Hildebrandt PK, Maronpot RR, Prater DA, Riley JH, Seaman WJ, Thake DC. 2004. Best practices guideline: toxicologic histopathology. Toxicol.Pathol. 32: 126-131.

EC (European Commission) Scientific Committee on Food. 2001. Guidance on submissions for safety evaluation of sources of nutrients or other ingredients proposed for use in the manufacture of foods. Available at: http://ec.europa.eu/food/fs/sc/scf/out100_en.pdf

EFSA (European Food Safety Authority), 2005. Opinion of the Scientific Panel on Animal Health and Welfare on a request from the Commission related to "Aspects of the biology and welfare of animals used for experimental and other scientific purposes". The EFSA Journal, 292, 1-46 http://www.efsa.europa.eu/en/efsajournal/doc/292.pdf

EFSA (European Food Safety Authority), 2009. Opinion of the Scientific Committee on a request from EFSA on existing approaches incorporating replacement, reduction and refinement of animal testing: applicability in food and feed risk assessment. The EFSA Journal, 1052, 1-77.

EFSA (European Food Safety Authority), 2011. Scientific opinion on Statistical significance and biological biological relevance. EFSA Journal 9(9):2372 [17 pp.].

EFSA GMO Panel Working Group on Animal Feeding Trials, 2008. Safety and nutritional assessment of GM plants and derived foodand feed: The role of animal feeding trials. Food and Chemical Toxicology 46:S2-S70.

EFSA Panel on Genetically Modified Organisms (GMO), 2010. Scientific Opinion on Statistical considerations for the safety evaluation of GMOs. EFSA Journal 8(1):1250, [59 pp.]

EFSA Panel on Genetically Modified Organisms (GMO), 2011a. Scientific Opinion on Guidance for risk assessment of food and feed from genetically modified plants. EFSA Journal 9(5): 2150. [37 pp.]

EFSA Panel on Genetically Modified Organisms (GMO), 2011b. Guidance document on Selection of Comparators for the Risk Assessment of GM Plants. EFSA Journal 9(5):2149. [21 pp.]

Ellis PD. 2010. The essential guide to effect sizes. Cambridge: Cambridge University Press, 2010: 3-173.

FDA (Food and Drug Administration) Redbook, 2000. Guidance for Industry and Other Stakeholders, Toxicological Principles for the Safety Assessment of Food Ingredients. July 2000; Revised July 2007. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Food Safety and Applied Nutrition. http://www.cfsan.fda.gov/guidance.html.

FELASA, 1996. Recommendations for the health monitoring of mouse, rat, hamster, gerbil, guinea pig and rabbit experimental units. Report of the Federation of European Laboratory Animal Science

Associations (FELASA) Working Group on Animal Health accepted by the FELASA Board of Management, November 1995. Laboratory Animals, 30, 193-208.

Festing MFW and Lutz C, 2010. Introduction to laboratory animal genetics. In: Hubrecht R, Kirkwood J, editors. The care and Management of Laboratory and Other Research Animals. 8th. ed. Oxford, Ames: Wiley-Blackwell, pp. 37-60.

Festing MFW and Lutz C, 2011. Laboratory Animal Genetics and Genetic Quality Control. In: Hau J, Schapiro SJ, editors. Handbook of Laboratory Animals Science. 3rd. ed. CRC Press, 209-50.

Fisher RA, 1960. The design of experiments. New York: Hafner Publishing Company, Inc, 1-248.

Knudsen I and Poulsen M, 2007. Comparative Safety Testing of Genetically Modified Foods in a 90-Day Rat Feeding Study Design Allowing the Distinction between Primary and Secondary Effects of the New Genetic Event. Regulatory Toxicology and Pharmacology, 49, 53-62.

Lapiz MD, Fulford A, Muchimapura S, Mason R, Parker T, Marsden CA. 2001. Influence of postweaning social isolation in the rat on brain development, conditioned behaviour and neurotransmission. Ross.Fiziol.Zh.Im I.M.Sechenova, **87:** 730-751

Leshem M and Sherman M, 2006. Troubles shared are troubles halved: stress in rats is reduced in proportion to social propinquity. Physiol Behav, 89(3):399-401.

Mead R, 1988. The design of experiments. Cambridge, New York: Cambridge University Press, pp. 1-620.

Montgomery DC, 1984. Design and Analysis of Experiments. Ed. John Wiley & Sons, Inc. New York.

Myhrer T. 1998. Adverse psychological impact, glutamatergic dysfunction, and risk factors for Alzheimer's disease. *Neurosci.Biobehav.Rev.*; **23:** 131-139.

Nakagawa S and Cuthill IC, 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol Rev Camb Philos Soc, 82:591-605.

NRC (National Research Council), 2005. Nutrient requirements of laboratory animals. Fourth revised edition, 1995.

OECD. 1998a. Guideline for the Testing of Chemicals – Repeated Dose 90-day Oral Toxicity Study in Rodents, 408. http://browse.oecdbookshop.org/oecd/pdfs/free/9740801e.pdf

OECD. 1998b. Principles of Good Laboratory Practice (GLP) and Compliance Monitoring. http://www.oecd.org/document/63/0,2340,en_2649_34381_2346175_1_1_1_1,00.html

OECD. 2008. Guideline for the Testing of Chemicals – Repeated Dose 28-day Oral Toxicity Study in Rodents, 407. http://browse.oecdbookshop.org/oecd/pdfs/free/9740701e.pdf

Poulsen M, Schrøder M, Wilcks A, Kroghsbo S, Lindecrona RH, Miller A, Frenzel T, Danier J, Rychlik M, Shu Q, Emami K, Taylor M, Gatehouse A, Engel KH, Knudsen I, 2007. Safety testing of GM-rice expressing PHA-E lectin using a new animal test design. Food/feed Chem. Toxicol. 45, 364-377.

Rao GN and Knapka JJ, 1987. Contaminant and Nutrient Concentrations of Natural Ingredient Rat and Mouse Diet Used in Chemical Toxicology Studies. Fundam Appl. Toxicol. 9, 329–338.

Stevens KA and Russel RJ, 2007. Chapter 10: Nutrition. The mouse in the biomedical research. Eds JG Fox, S Barthold, M Davisson, CE Newcomer, FW Quimby, A Smith. II Edition Academic Press.

TASCC (Trade Assurance Scheme for Combinable Crops), 2010. Code of Practice for the Storage of Combinable Crops and Animal Feeds. Effective from July 1st 2010. http://www.agindustries.org.uk/document.aspx?fn=load&media_id=3734&publicationId=2150

Verhagen H, Aruoma OI, van Delft JHM., Dragsted LO, Ferguson LR, Knasmüller S, Pool-Zobel BL, Poulsen HE, Williamson G, Yannai S, 2003. Editorial - The 10 basic requirements for a scientific

paper reporting antioxidant, antimutagenic or anticarcinogenic potential of test substances in in vitro experiments and animal studies in vivo. Food and Chemical Toxicology 41: 603-610.

Westenbroek C, Ter Horst GJ, Roos MH, Kuipers SD, Trentani A, den Boer JA, 2003. Gender-specific effects of social housing in rats after chronic mild stress exposure. Prog. Neuropsychopharmacol. Biol. Psychiatry 27(1):21-30.

GLOSSARY AND ABBREVIATIONS

Term	Description
Dose (OECD)	The amount of test substance administered. Dose is expressed as weight (g, mg) or as weight of test substance per unit body weight of test animal (e.g., mg/kg bw), or as constant dietary concentrations (ppm).
Dosage (OECD)	A general term comprising of dose, its frequency and the duration of dosing.
ExpU	Experimental unit(s). The smallest division of the experimental material such that any two ExpU can receive different treatments.
GLP	Good Laboratory Practice
NOAEL	No observed adverse effect level. The highest dose level where no adverse treatment-related findings are observed.
SAP	Statistical Analysis Plan
SD	Standard Deviation
SPF	Specific Pathogen Free. Status for laboratory animals that are guaranteed free of particular pathogens.
TG	Technical Guidance