OECD/OCDE 474
Adopted:
29 July 2016

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Mammalian Erythrocyte Micronucleus Test

INTRODUCTION

1. The OECD Guidelines for the Testing of Chemicals are periodically reviewed in the light of scientific progress, changing regulatory needs and animal welfare considerations. The original Test Guideline 474 was adopted in 1983. In 1997, a revised version was adopted, based on scientific progress made to that date. This modified version of the Test Guideline reflects scientific knowledge from more than thirty years of experience with this assay and the interpretation of the data, and in particular the advances in automated scoring technologies and the potential for integrating or combining this test with other general toxicity or genotoxicity studies. This Test Guideline is part of a series of Test Guidelines on genetic toxicology. A document that provides succinct information on genetic toxicology testing and an overview of the recent changes that were made to these Test Guidelines has been developed (1).

2. The mammalian *in vivo* micronucleus test is especially relevant for assessing genotoxicity because, although they may vary among species, factors of *in vivo* metabolism, pharmacokinetics and DNA repair processes are active and contribute to the responses. An *in vivo* assay is also useful for further investigation of genotoxicity detected by an *in vitro* system.

3. The mammalian *in vivo* micronucleus test is used for the detection of damage induced by the test chemical to the chromosomes or the mitotic apparatus of erythroblasts. The test evaluates micronucleus formation in erythrocytes sampled either in the bone marrow or peripheral blood cells of animals, usually rodents.

4. The purpose of the micronucleus test is to identify substances that cause cytogenetic damage which results in the formation of micronuclei containing either lagging chromosome fragments or whole chromosomes.

5. When a bone marrow erythroblast develops into an immature erythrocyte (sometimes also referred to as a polychromatic erythrocyte or reticulocyte), the main nucleus is extruded; any micronucleus that has been formed may remain behind in the cytoplasm. Visualisation or detection of micronuclei is facilitated in these cells because they lack a main nucleus. An increase in the frequency of micronucleated

This Guideline was adopted by the OECD Council by written procedure on 29 July 2016 [C(2016)103].

immature erythrocytes in treated animals is an indication of induced structural or numerical chromosomal aberrations.

6. Newly formed micronucleated erythrocytes are identified and quantitated by staining followed by either visual scoring using a microscope, or by automated analysis. Counting sufficient immature erythrocytes in the peripheral blood or bone marrow of adult animals is greatly facilitated by using an automated scoring platform. Such platforms are acceptable alternatives to manual evaluation (2). Comparative studies have shown that such methods, using appropriate calibration standards, can provide better inter- and intra-laboratory reproducibility and sensitivity than manual microscopic scoring (3) (4). Automated systems that can measure micronucleated erythrocyte frequencies include, but are not limited to, flow cytometers (5), image analysis platforms (6) (7), and laser scanning cytometers (8).

7. Although not normally done as part of the test, chromosome fragments can be distinguished from whole chromosomes by a number of criteria. These include identification of the presence or absence of a kinetochore or centromeric DNA, both of which are characteristic of intact chromosomes. The absence of kinetochore or centromeric DNA indicates that the micronucleus contains only fragments of chromosomes, while the presence is indicative of chromosome loss.

8. Definitions of terminology used are set out in Annex 1.

INITIAL CONSIDERATIONS

9. The bone marrow of young adult rodents is the target tissue for genetic damage in this test since erythrocytes are produced in this tissue. The measurement of micronuclei in immature erythrocytes in peripheral blood is acceptable in other mammalian species for which adequate sensitivity to detect substances that cause structural or numerical chromosomal aberrations in these cells has been demonstrated (by induction of micronuclei in immature erythrocytes) and scientific justification is provided. The frequency of micronucleated immature erythrocytes is the principal endpoint. The frequency of mature erythrocytes that contain micronuclei in the peripheral blood also can be used as an endpoint in species without strong splenic selection against micronucleated cells and when animals are treated continuously for a period that exceeds the lifespan of the erythrocyte in the species used (e.g. 4 weeks or more in the mouse).

10. If there is evidence that the test substance(s), or its metabolite(s), will not reach the target tissue, it may not be appropriate to use this test.

11. Before use of the Test Guideline on a mixture for generating data for an intended regulatory purpose, it should be considered whether, and if so why, it may provide adequate results for that purpose. Such considerations are not needed, when there is a regulatory requirement for testing of the mixture.

PRINCIPLE OF THE TEST METHOD

12. Animals are exposed to the test chemical by an appropriate route. If bone marrow is used, the animals are humanely euthanised at an appropriate time(s) after treatment, the bone marrow is extracted, and preparations are made and stained (9) (10) (11) (12) (13) (14) (15). When peripheral blood is used, the blood is collected at an appropriate time(s) after treatment and preparations are made and stained (12) (16) (17) (18). When treatment is administered acutely, it is important to select bone marrow or blood harvest

times at which the treatment-related induction of micronucleated immature erythrocytes can be detected. In the case of peripheral blood sampling, enough time must also have elapsed for these events to appear in circulating blood. Preparations are analysed for the presence of micronuclei, either by visualisation using a microscope, image analysis, flow cytometry, or laser scanning cytometry.

VERIFICATION OF LABORATORY PROFICIENCY

Proficiency Investigations

13. In order to establish sufficient experience with the conduct of the assay prior to using it for routine testing, the laboratory should have demonstrated the ability to reproduce expected results from published data (17) (19) (20) (21) (22) for micronucleus frequencies with a minimum of two positive control substances (including weak responses induced by low doses of positive controls), such as those listed in Table 1 and with compatible vehicle/solvent controls (see paragraph 26). These experiments should use doses that give reproducible and dose-related increases and demonstrate the sensitivity and dynamic range of the test system in the tissue of interest (bone marrow or peripheral blood) and using the scoring method to be employed within the laboratory. This requirement is not applicable to laboratories that have experience, i.e. that have a historical database available as defined in paragraphs 14-18.

Historical Control Data

14. During the course of the proficiency investigations, the laboratory should establish:

- A historical positive control range and distribution, and
- A historical negative control range and distribution.

15. When first acquiring data for a historical negative control distribution, concurrent negative controls should be consistent with published control data, where they exist. As more experimental data are added to the historical control distribution, concurrent negative controls should ideally be within the 95% control limits of that distribution. The laboratory's historical negative control database should be statistically robust to ensure the ability of the laboratory to assess the distribution of their negative control data. The literature suggests that a minimum of 10 experiments may be necessary but would preferably consist of at least 20 experiments conducted under comparable experimental conditions. Laboratories should use quality control methods, such as control charts (e.g. C-charts or X-bar charts (23)), to identify how variable their data are, and to show that the methodology is 'under control' in their laboratory. Further recommendations on how to build and use the historical data (i.e. criteria for inclusion and exclusion of data in historical data and the acceptability criteria for a given experiment) can be found in the literature (24).

16. Where the laboratory does not complete a sufficient number of experiments to establish a statistically robust negative control distribution (see paragraph 15) during the proficiency investigations (described in paragraph 13), it is acceptable that the distribution can be built during the first routine tests. This approach should follow the recommendations set out in the literature (24) and the negative control results obtained in these experiments should remain consistent with published negative control data.

17. Any changes to the experimental protocol should be considered in terms of their impact on the resulting data remaining consistent with the laboratory's existing historical control database. Only major inconsistencies should result in the establishment of a new historical control database where expert

judgement determines that it differs from the previous distribution (see paragraph 15). During the re-establishment, a full negative control database may not be needed to permit the conduct of an actual test, provided that the laboratory can demonstrate that their concurrent negative control values remain either consistent with their previous database or with the corresponding published data.

18. Negative control data should consist of the incidence of micronucleated immature erythrocytes in each animal. Concurrent negative controls should ideally be within the 95% control limits of the distribution of the laboratory's historical negative control database. Where concurrent negative control data fall outside the 95% control limits, they may be acceptable for inclusion in the historical control distribution as long as these data are not extreme outliers and there is evidence that the test system is 'under control' (see paragraph 15) and no evidence of technical or human failure.

DESCRIPTION OF THE METHOD

Preparations

Selection of animal species

19. Commonly used laboratory strains of healthy young adult animals should be employed. Mice, rats, or another appropriate mammalian species may be used. When peripheral blood is used, it must be established that splenic removal of micronucleated cells from the circulation does not compromise the detection of induced micronuclei in the species selected. This has been clearly demonstrated for mouse and rat peripheral blood (2). The scientific justification for using species other than rats and mice should be provided in the report. If species other than rodents are used, it is recommended that the measurement of induced micronuclei be integrated into another appropriate toxicity test.

Animal housing and feeding conditions

20. For rodents, the temperature in the animal room should be 22°C (±3°C). Although the relative humidity ideally should be 50-60%, it should be at least 40% and preferably not exceed 70% other than during room cleaning. Lighting should be artificial, the sequence being 12 hours light, 12 hours dark. For feeding, conventional laboratory diets may be used with an unlimited supply of drinking water. The choice of diet may be influenced by the need to ensure a suitable admixture of a test chemical when administered by this route. Rodents should be housed in small groups (no more than five per cage) of the same sex and treatment group if no aggressive behaviour is expected, preferably in solid floor cages with appropriate environmental enrichment. Animals may be housed individually only if scientifically justified.

Preparation of the animals

21. Healthy young adult animals (for rodents, ideally 6-10 weeks old at start of treatment, though slightly older animals are also acceptable) are normally used, and are randomly assigned to the control and treatment groups. The individual animals are identified uniquely using a humane, minimally invasive method (e.g. by ringing, tagging, micro-chipping or biometric identification, but not ear or toe clipping) and acclimated to the laboratory conditions for at least five days. Cages should be arranged in such a way that possible effects due to cage placement are minimised. Cross contamination by the positive control and the test chemical should be avoided. 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.

Preparation of doses

22. Solid test chemicals should be dissolved or suspended in appropriate solvents or vehicles or admixed in diet or drinking water prior to dosing the animals. Liquid test chemicals may be dosed directly or diluted prior to dosing. For inhalation exposures, test chemicals can be administered as a gas, vapour, or a solid/liquid aerosol, depending on their physicochemical properties. Fresh preparations of the test chemical should be employed unless stability data demonstrate the acceptability of storage and define the appropriate storage conditions.

Test Conditions

Solvent/vehicle

23. The solvent/vehicle should not produce toxic effects at the dose levels used, and should not be capable of chemical reaction with the test substances. If other than well-known solvents/vehicles are used, their inclusion should be supported with reference data indicating their compatibility. It is recommended that wherever possible, the use of an aqueous solvent/vehicle should be considered first. Examples of commonly used compatible solvents/vehicles include water, physiological saline, methylcellulose solution, carboxymethyl cellulose sodium salt solution, olive oil and corn oil. In the absence of historical or published control data showing that no micronuclei and other deleterious effects are induced by a chosen atypical solvent/vehicle, an initial study should be conducted in order to establish the acceptability of the solvent/vehicle control.

Controls

Positive controls

24. A group of animals treated with a positive control substance should normally be included with each test. This may be waived when the testing laboratory has demonstrated proficiency in the conduct of the test and has established a historical positive control range. When a concurrent positive control group is not included, scoring controls (fixed and unstained slides or cell suspension samples, as appropriate for the method of scoring) should be included in each experiment. These can be obtained by including within the scoring of the study appropriate reference samples that have been obtained and stored from a separate positive control experiment conducted periodically (e.g. every 6-18 months); for example, during proficiency testing and on a regular basis thereafter, where necessary.

25. Positive control substances should reliably produce a detectable increase in micronucleus frequency over the spontaneous level. When employing manual scoring by microscopy, positive control doses should be chosen so that the effects are clear but do not immediately reveal the identity of the coded samples to the scorer. It is acceptable that the positive control be administered by a route different from the test chemical, using a different treatment schedule, and for sampling to occur only at a single time point. In addition, the use of chemical class-related positive control substances may be considered, when appropriate. Examples of positive control substances are included in Table 1.

Table 1. Examples of positive control substances.

Substances and CASRN
Ethyl methanesulphonate [CASRN 62-50-0]
Methyl methanesulphonate [CASRN 66-27-3]
Ethyl nitrosourea [CASRN 759-73-9]
Mitomycin C [CASRN 50-07-7]
Cyclophosphamide (monohydrate) [CASRN 50-18-0 (CASRN 6055-19-2)]
Triethylenemelamine [CASRN 51-18-3]
Colchicine [CASRN 64-86-8] or Vinblastine [CASRN 865-21-4] – as aneugens

Negative controls

26. Negative control group animals should be included at every sampling time and otherwise handled in the same way as the treatment groups, except for not receiving treatment with the test chemical. If a solvent/vehicle is used in administering the test chemical, the control group should receive this solvent/vehicle. However, if consistent inter-animal variability and frequencies of cells with micronuclei are demonstrated by historical negative control data at each sampling time for the testing laboratory, only a single sampling for the negative control may be necessary. Where a single sampling is used for negative controls, it should be the first sampling time used in the study.

27. If peripheral blood is used, a pre-treatment sample is acceptable instead of a concurrent negative control for short-term studies when the resulting data are consistent with the historical control database for the testing laboratory. It has been shown for rats that pre-treatment sampling of small volumes (e.g. below 100 µL/day) has minimal impact on micronucleus background frequency (25).

PROCEDURE

Number and sex of animals

28. In general, the micronucleus response is similar between male and female animals and, therefore, most studies could be performed in either sex (26). Data demonstrating relevant differences between males and females (e.g. differences in systemic toxicity, metabolism, bioavailability, bone marrow toxicity, etc. including e.g. in a range-finding study) would encourage the use of both sexes. In this case, it may be appropriate to perform a study in both sexes, e.g. as part of a repeated dose toxicity study. It might be appropriate to use the factorial design in case both sexes are used. Details on how to analyse the data using this design are given in Annex 2.

29. Group sizes at study initiation should be established with the aim of providing a minimum of 5 analysable animals of one sex, or of each sex if both are used, per group. Where human exposure to chemicals may be sex-specific, as for example with some pharmaceuticals, the test should be performed

with the appropriate sex. As a guide to maximum typical animal requirements, a study in bone marrow conducted according the parameters established in paragraph 37 with three dose groups and concurrent negative and positive controls (each group composed of five animals of a single sex) would require between 25 and 35 animals.

Dose levels

30. If a preliminary range-finding study is performed because there are no suitable data already available to aid in dose selection, it should be performed in the same laboratory, using the same species, strain, sex, and treatment regimen to be used in the main study (27). The study should aim to identify the maximum tolerated dose (MTD), defined as the highest dose that will be tolerated without evidence of study-limiting toxicity, relative to the duration of the study period (for example, by inducing body weight depression or hematopoietic system cytotoxicity, but not death or evidence of pain, suffering or distress necessitating humane euthanasia (28)).

31. The highest dose may also be defined as a dose that produces toxicity in the bone marrow (e.g. a reduction in the proportion of immature erythrocytes among total erythrocytes in the bone marrow or peripheral blood of more than 50%, but to not less than 20% of the control value). However, when analysing CD71-positive cells in peripheral blood circulation (i.e., by flow cytometry), this very young fraction of immature erythrocytes responds to toxic challenges more quickly than the larger RNA-positive cohort of immature erythrocytes. Therefore, higher apparent toxicity may be evident with acute exposure designs examining the CD71-positive immature erythrocyte fraction as compared to those that identify immature erythrocytes based on RNA content. For this reason, when experiments utilise five or fewer days of treatment, the highest dose level for test chemicals causing toxicity may be defined as the dose that causes a statistically significant reduction in the proportion of CD71-positive immature erythrocytes among total erythrocytes but not to less than 5% of the control value (29).

32. Substances that exhibit saturation of toxicokinetic properties, or induce detoxification processes that may lead to a decrease in exposure after long-term administration may be exceptions to the dose-setting criteria and should be evaluated on a case-by-case basis.

33. In order to obtain dose response information, a complete study should include a negative control group and a minimum of three dose levels generally separated by a factor of 2, but by no greater than 4. If the test chemical does not produce toxicity in a range-finding study or based on existing data, the highest dose for an administration period of 14 days or more should be 1000 mg/kg body weight/day, or for administration periods of less than 14 days, 2000 mg/kg/body weight/day. However, if the test chemical does cause toxicity, the MTD should be the highest dose administered and the dose levels used should preferably cover a range from the maximum to a dose producing little or no toxicity. When target tissue (bone marrow) toxicity is observed at all dose levels tested, further study at non-toxic doses is advisable. Studies intending to more fully characterise the quantitative dose-response information may require additional dose groups. For certain types of test chemicals (e.g. human pharmaceuticals) covered by specific requirements, these limits may vary.

Limit test

34. If dose range-finding experiments, or existing data from related animal strains, indicate that a treatment regime of at least the limit dose (described below) produces no observable toxic effects, (including no depression of bone marrow proliferation or other evidence of target tissue cytotoxicity), and if genotoxicity would not be expected based upon *in vitro* genotoxicity studies or data from structurally

related substances, then a full study using three dose levels may not be considered necessary, provided it has been demonstrated that the test chemical(s) reach(es) the target tissue (bone marrow). In such cases, a single dose level, at the limit dose, may be sufficient. When administration occurs for 14 days or more, the limit dose is 1000 mg/kg body weight/day. For administration periods of less than 14 days, the limit dose is 2000 mg/kg/body weight/day.

Administration of doses

35. The anticipated route of human exposure should be considered when designing an assay. Therefore, routes of exposure such as dietary, drinking water, topical subcutaneous, intravenous, oral (by gavage), inhalation, intratracheal, or implantation may be chosen as justified. In any case, the route should be chosen to ensure adequate exposure of the target tissue(s). Intraperitoneal injection is generally not recommended since it is not an intended route of human exposure, and should only be used with specific scientific justification. If the test chemical is admixed in diet or drinking water, especially in case of single dosing, care should be taken that the delay between food and water consumption and sampling should be sufficient to allow detection of the effects (see paragraph 37). The maximum volume of liquid that can be administered by gavage or injection at one time depends on the size of the test animal. The volume should not normally exceed 1 mL/100 g body weight except in the case of aqueous solutions where a maximum of 2 mL/100 g may be used. The use of volumes greater than this should be justified. Except for irritating or corrosive test chemicals, which will normally produce exacerbated effects at higher concentrations, variability in test volume should be minimised by adjusting the concentration to ensure administration of a constant volume in relation to body weight at all dose levels.

Treatment schedule

36. Preferably, 2 or more treatments are performed, administered at 24-hour intervals, especially when integrating this test into other toxicity studies. In the alternative, single treatments can be administered, if scientifically justified (e.g. test chemicals known to block cell cycle). Test chemicals also may be administered as a split dose, i.e., two or more treatments on the same day separated by no more than 2-3 hours, to facilitate administering a large volume. Under these circumstances, or when administering the test chemical by inhalation, the sampling time should be scheduled based on the time of the last dosing or the end of exposure.

37. The test may be performed in mice or rats in one of three ways:

a. Animals are treated with the test chemical once. Samples of bone marrow are taken at least twice (from independent groups of animals), starting not earlier than 24 hours after treatment, but not extending beyond 48 hours after treatment with appropriate interval(s) between samples, unless a test substance is known to have an exceptionally long half-life. The use of sampling times earlier than 24 hours after treatment should be justified. Samples of peripheral blood are taken at least twice (from the same group of animals), starting not earlier than 36 hours after treatment, with appropriate intervals following the first sample, but not extending beyond 72 hours. At the first sampling time, all dose groups should be treated and samples collected for analysis; however, at the later sampling time(s), only the highest dose needs to be administered. When a positive response is detected at one sampling time, additional sampling is not required unless quantitative dose-response information is needed. The described harvest times are a consequence of the kinetics of appearance and disappearance of the micronuclei in these 2 tissue compartments.

b. If 2 daily treatments are used (e.g. two treatments at 24 hour intervals), samples should be collected once between 18 and 24 hours following the final treatment for the bone marrow or once between 36 and 48 hours following the final treatment for peripheral blood (30). The described harvest times are a consequence of the kinetics of appearance and disappearance of the micronuclei in these 2 tissue compartments.

c. If three or more daily treatments are used (e.g. three or more treatments at approximately 24 hour intervals), bone marrow samples should be collected no later than 24 hours after the last treatment and peripheral blood should be collected no later than 40 hours after the last treatment (31). This treatment option accommodates combination of the comet assay (e.g. sampling 2-6 hours after the last treatment) with the micronucleus test, and integration of the micronucleus test with repeated-dose toxicity studies. Accumulated data suggested that micronucleus induction can be observed over these wider timeframes when 3 or more administrations have occurred (15).

38. Other dosing or sampling regimens may be used when relevant and scientifically justified, and to facilitate integration with other toxicity tests.

Observations

39. General clinical observations of the test animals should be made and clinical signs recorded at least once a day, preferably at the same time(s) each day and considering the peak period of anticipated effects after dosing. At least twice daily during the dosing period, all animals should be observed for morbidity and mortality. All animals should be weighed at study initiation, at least once a week during repeated dose studies, and at euthanasia. In studies of at least one-week duration, measurements of food consumption should be made at least weekly. If the test chemical is administered via the drinking water, water consumption should be measured at each change of water and at least weekly. Animals exhibiting non-lethal indicators of excessive toxicity should be humanely euthanised prior to completion of the test period (28). Under certain circumstances, animal body temperature could be monitored, since treatment-induced hyper- and hypothermia have been implicated in producing spurious results (32) (33) (34).

Target tissue exposure

40. A blood sample should be taken at appropriate time(s) in order to permit investigation of the plasma levels of the test substances for the purposes of demonstrating that exposure of the bone marrow occurred, where warranted and where other exposure data do not exist (see paragraph 48).

Bone marrow / blood preparation

41. Bone marrow cells are usually obtained from the femurs or tibias of the animals immediately following humane euthanasia. Commonly, cells are removed, prepared and stained using established methods. Small volumes of peripheral blood can be obtained, according to adequate animal welfare standards, either using a method that permits survival of the test animal, such as bleeding from the tail vein or other appropriate blood vessel, or by cardiac puncture or sampling from a large vessel at animal euthanasia. For both bone marrow or peripheral blood-derived erythrocytes, depending on the method of analysis, cells may be immediately stained supravitally (16) (17) (18), smear preparations are made and then stained for microscopy, or fixed and stained appropriately for flow cytometric analysis. The use of a DNA specific stain [e.g. acridine orange (35) or Hoechst 33258 plus pyronin-Y (36)] can eliminate some of the artifacts associated with using a non-DNA specific stain. This advantage does not preclude the use

of conventional stains (e.g. Giemsa for microscopic analysis). Additional systems [e.g. cellulose columns to remove nucleated cells (37) (38)] also can be used provided that these systems have been demonstrated to be compatible with sample preparation in the laboratory.

42. Where these methods are applicable, anti-kinetochore antibodies (39), FISH with pancentromeric DNA probes (40), or primed *in situ* labelling with pancentromere-specific primers, together with appropriate DNA counterstaining (41), can be used to identify the nature of the micronuclei (chromosome/chromosomal fragment) in order to determine whether the mechanism of micronucleus induction is due to clastogenic and/or aneugenic activity. Other methods for differentiation between clastogens and aneugens may be used if they have been shown to be effective.

Analysis (manual and automated)

43. All slides or samples for analysis, including those of positive and negative controls, should be independently coded before any type of analysis and should be randomised so the manual scorer is unaware of the treatment condition; such coding is not necessary when using automated scoring systems which do not rely on visual inspection and cannot be affected by operator bias. The proportion of immature among total (immature + mature) erythrocytes is determined for each animal by counting a total of at least 500 erythrocytes for bone marrow and 2000 erythrocytes for peripheral blood (42). At least 4000 immature erythrocytes per animal should be scored for the incidence of micronucleated immature erythrocytes (43). If the historical negative control database indicates the mean background micronucleated immature erythrocyte frequency is <0.1% in the testing laboratory, consideration should be given to scoring additional cells. When analysing samples, the proportion of immature erythrocytes to total erythrocytes in treated animals should not be less than 20% of the vehicle/solvent control proportion when scoring by microscopy and not less than approximately 5% of the vehicle/solvent control proportion when scoring CD71+ immature erythrocytes by cytometric methods (see paragraph 31) (29). For example for a bone marrow assay scored by microscopy, if the control proportion of immature erythrocytes in the bone marrow is 50%, the upper limit of toxicity would be 10% immature erythrocytes.

44. Because the rat spleen sequesters and destroys micronucleated erythrocytes, to maintain high assay sensitivity when analysing rat peripheral blood, it is preferable to restrict the analysis of micronucleated immature erythrocytes to the youngest fraction. When using automated analysis methods, these most immature erythrocytes can be identified based on their high RNA content, or the high level of transferrin receptors (CD71+) expressed on their surface (31). However, direct comparison of different staining methods has shown that satisfactory results can be obtained with various methods, including conventional acridine orange staining (3) (4).

DATA AND REPORTING

Treatment of Results

45. Individual animal data should be presented in tabular form. The number of immature erythrocytes scored, the number of micronucleated immature erythrocytes, and the proportion of immature among total erythrocytes should be listed separately for each animal analysed. When mice are treated continuously for 4 weeks or more, the data on the number and proportion of micronucleated mature erythrocytes also should be given if collected. Data on animal toxicity and clinical signs should also be reported.

Acceptability Criteria

46. The following criteria determine the acceptability of the test:

a) The concurrent negative control data are considered acceptable for addition to the laboratory historical control database (see paragraphs 15-18).

b) The concurrent positive controls or scoring controls should induce responses that are compatible with those generated in the historical positive control database and produce a statistically significant increase compared with the concurrent negative control (see paragraphs 24-25).

c) The appropriate number of doses and cells has been analysed.

d) The criteria for the selection of highest dose are consistent with those described in paragraphs 30-33.

Evaluation and Interpretation of Results

47. Providing that all acceptability criteria are fulfilled, a test chemical is considered clearly positive if:

a) At least one of the treatment groups exhibits a statistically significant increase in the frequency of micronucleated immature erythrocytes compared with the concurrent negative control,

b) This increase is dose-related at least at one sampling time when evaluated with an appropriate trend test, and

c) Any of these results are outside the distribution of the historical negative control data (e.g. Poisson-based 95% control limits).

If only the highest dose is examined at a particular sampling time, a test chemical is considered clearly positive if there is a statistically significant increase compared with the concurrent negative control and the results are outside the distribution of the historical negative control data (e.g. Poisson-based 95% control limits). Recommendations for the most appropriate statistical methods can be found in the literature (44) (45) (46) (47). When conducting a dose-response analysis, at least three treated dose groups should be analysed. Statistical tests should use the animal as the experimental unit. Positive results in the micronucleus test indicate that a test chemical induces micronuclei, which are the result of chromosomal damage or damage to the mitotic apparatus in the erythroblasts of the test species. In the case where a test was performed to detect centromeres within micronuclei, a test chemical that produces centromere-containing micronuclei (centromeric DNA or kinetochore, indicative of whole chromosome loss) is evidence that the test chemical is an aneugen.

48. Providing that all acceptability criteria are fulfilled, a test chemical is considered clearly negative if, in all experimental conditions examined:

a) None of the treatment groups exhibits a statistically significant increase in the frequency of micronucleated immature erythrocytes compared with the concurrent negative control,

b) There is no dose-related increase at any sampling time when evaluated by an appropriate trend test,

c) All results are inside the distribution of the historical negative control data (e.g. Poisson-based 95% control limits), and

d) Bone marrow exposure to the test substance(s) occurred.

Recommendations for the most appropriate statistical methods can be found in the literature (44) (45) (46) (47). Evidence of exposure of the bone marrow to a test substance may include a depression of the immature to mature erythrocyte ratio or measurement of the plasma or blood levels of the test substance. In case of intravenous administration, evidence of exposure is not needed. Alternatively, ADME data, obtained in an independent study using the same route and same species can be used to demonstrate bone marrow exposure. Negative results indicate that, under the test conditions, the test chemical does not produce micronuclei in the immature erythrocytes of the test species.

49. There is no requirement for verification of a clear positive or clear negative response.

50. In cases where the response is not clearly negative or positive and in order to assist in establishing the biological relevance of a result (e.g. a weak or borderline increase), the data should be evaluated by expert judgement and/or further investigations of the existing experiments completed. In some cases, analysing more cells or performing a repeat experiment using modified experimental conditions could be useful.

51. In rare cases, even after further investigations, the data will preclude making a conclusion that the test chemical produces either positive or negative results, and the study will therefore be concluded as equivocal.

Test Report

52. The test report should include the following information:

Summary

Test chemical:

- source, lot number, limit date for use, if available;
- stability of the test chemical, if known.

Mono-constituent substance:

- physical appearance, water solubility, and additional relevant physicochemical properties;
- chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities as appropriate and practically feasible, etc.

Multi-constituent substance, UVBCs and mixtures:

- characterised as far as possible by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents.

Test chemical preparation:

- justification for choice of vehicle;
- solubility and stability of the test chemical in the solvent/vehicle, if known;
- preparation of dietary, drinking water or inhalation formulations;
- analytical determinations on formulations (e.g. stability, homogeneity, nominal concentrations), when conducted.

Test animals:

- species/strain used and justification for use;
- number, age and sex of animals;
- source, housing conditions, diet, etc.;
- method for uniquely identifying the animals;
- for short term studies: individual weight of the animals at the start and end of the test; for studies longer than one week: individual body weights during the study and food consumption. Body weight range, mean and standard deviation for each group should be included.

Test conditions:

- positive and negative (vehicle/solvent) control data;
- data from range-finding study, if conducted;
- rationale for dose level selection;
- details of test chemical preparation;
- details of the administration of the test chemical;
- rationale for route and duration of administration;
- methods for verifying that the test substance(s) reached the general circulation or target tissue;
- actual dose (mg/kg body weight/day) calculated from diet/drinking water test chemical concentration (ppm) and consumption, if applicable;
- details of food and water quality;
- method of euthanasia;
- method of analgesia (where used);
- detailed description of treatment and sampling schedules and justifications for the choices;
- methods of slide preparation;
- procedures for isolating and preserving samples;
- methods for measurement of toxicity;
- criteria for scoring micronucleated immature erythrocytes;
- number of cells analysed per animal in determining the frequency of micronucleated immature erythrocytes and for determining the proportion of immature to mature erythrocytes;
- criteria for acceptability of the study;
- methods, such as use of anti-kinetochore antibodies or centromere-specific DNA probes, to characterise whether micronuclei contain whole or fragmented chromosomes, if applicable.

Results:

- animal condition prior to and throughout the test period, including signs of toxicity;

- proportion of immature erythrocytes among total erythrocytes;
- number of micronucleated immature erythrocytes, given separately for each animal;
- mean ± standard deviation of micronucleated immature erythrocytes per group;
- dose-response relationship, where possible;
- statistical analyses and methods applied;
- concurrent negative and positive control data with ranges, means and standard deviations;
- historical negative and positive control data with ranges, means, standard deviations and 95% control limits for the distribution, as well as the time period covered and the number of data points;
- data supporting exposure of the bone marrow occurred;
- characterisation data indicating whether micronuclei contain whole or fragmented chromosomes, if applicable;
- criteria for a positive or negative response that are met.

Discussion of the results.

Conclusion.

References.

LITERATURE

(1) OECD (2016). Overview of the set of OECD Genetic Toxicology Test Guidelines and updates performed in 2014-2015. ENV Publications. Series on Testing and Assessment, No. 234, OECD, Paris.

(2) Hayashi, M. et al. (2007), In vivo erythrocyte micronucleus assay III. Validation and regulatory acceptance of automated scoring and the use of rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 627/1, pp. 10-30.

(3) MacGregor, J.T. et al. (2006), Flow cytometric analysis of micronuclei in peripheral blood reticulocytes: II. An efficient method of monitoring chromosomal damage in the rat, *Toxicology Sciences*, Vol. 94/1, pp. 92-107.

(4) Dertinger, S.D. et al. (2006), Flow cytometric analysis of micronuclei in peripheral blood reticulocytes: I. Intra- and interlaboratory comparison with microscopic scoring, *Toxicological Sciences*, Vol. 94/1, pp. 83-91.

(5) Dertinger, S.D. et al. (2011), Flow cytometric scoring of micronucleated erythrocytes: an efficient platform for assessing in vivo cytogenetic damage, *Mutagenesis*, Vol. 26/1, pp. 139-145.

(6) Parton, J.W., W.P. Hoffman, M.L. Garriott (1996), Validation of an automated image analysis micronucleus scoring system, *Mutation Research*, Vol. 370/1, pp. 65-73.

(7) Asano, N. et al. (1998), An automated new technique for scoring the rodent micronucleus assay: computerized image analysis of acridine orange supravitally stained peripheral blood cells, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, Vol. 404/1-2, pp. 149-154.

(8) Styles, J.A. et al. (2001), Automation of mouse micronucleus genotoxicity assay by laser scanning cytometry, *Cytometry*, Vol. 44/2, pp. 153-155.

(9) Heddle, J.A. (1973), A rapid in vivo test for chromosomal damage, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, Vol. 18/2, pp. 187-190.

(10) Schmid, W. (1975), The micronucleus test, *Mutation Research*, Vol. 31/1, pp. 9-15.

(11) Heddle, J.A. et al. (1983), The induction of micronuclei as a measure of genotoxicity. A report of the U.S. Environmental Protection Agency Gene-Tox Program, *Mutation Research/Reviews in Genetic Toxicology*, Vol. 123/1, pp. 61-118.

(12) Mavournin, K.H. et al. (1990), The in vivo micronucleus assay in mammalian bone marrow and peripheral blood. A report of the U.S. Environmental Protection Agency Gene-Tox Program, *Mutation Research/Reviews in Genetic Toxicology*, Vol. 239/1, pp. 29-80.

(13) MacGregor, J.T. et al. (1983), Micronuclei in circulating erythrocytes: a rapid screen for chromosomal damage during routine toxicity testing in mice, *Developments in Toxicology Environmental Science*, Vol. 11, pp. 555-558.

(14) MacGregor, J.T. et al. (1987), Guidelines for the conduct of micronucleus assays in mammalian bone marrow erythrocytes, *Mutation Research/Genetic Toxicology*, Vol. 189/2, pp. 103-112.

(15) MacGregor, J.T. et al. (1990), The in vivo erythrocyte micronucleus test: measurement at steady state increases assay efficiency and permits integration with toxicity studies, *Fundamental and Applied Toxicology*, Vol. 14/3, pp. 513-522.

(16) Hayashi, M. et al. (1990), The micronucleus assay with mouse peripheral blood reticulocytes using acridine orange-coated slides, *Mutation Research/Genetic Toxicology*, Vol. 245/4, pp. 245-249.

(17) CSGMT/JEMS.MMS – The Collaborative Study Group for the Micronucleus Test (1992), Micronucleus test with mouse peripheral blood erythrocytes by acridine orange supravital staining: the summary report of the 5th collaborative study, *Mutation Research/Genetic Toxicology*, Vol. 278/2-3, pp. 83-98.

(18) CSGMT/JEMS.MMS – The Mammalian Mutagenesis Study Group of the Environmental Mutagen Society of Japan (1995), Protocol recommended by the CSGMT/JEMS.MMS for the short-term mouse peripheral blood micronucleus test. The Collaborative Study Group for the Micronucleus Test (CSGMT) (CSGMT/JEMS.MMS, The Mammalian Mutagenesis Study Group of the Environmental Mutagen Society of Japan), *Mutagenesis*, Vol. 10/3, pp. 153-159.

(19) Salamone, M.F., K.H. Mavournin (1994), Bone marrow micronucleus assay: a review of the mouse stocks used and their published mean spontaneous micronucleus frequencies, *Environmental and Molecular Mutagenesis*, Vol. 23/4, pp. 239-273.

(20) Krishna, G., G. Urda, J. Paulissen (2000), Historical vehicle and positive control micronucleus data in mice and rats, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, Vol. 453/1, pp. 45-50.

(21) Hayes, J. et al. (2009), The rat bone marrow micronucleus test--study design and statistical power, *Mutagenesis*, Vol. 24/5, pp. 419-424.

(22) Wakata, A. et al. (1998), Evaluation of the rat micronucleus test with bone marrow and peripheral blood: summary of the 9th collaborative study by CSGMT/JEMS. MMS. Collaborative Study Group for the Micronucleus Test. Environmental Mutagen Society of Japan. Mammalian Mutagenicity Study Group, *Environmental and Molecular Mutagenesis*, Vol. 32/1, pp. 84-100.

(23) Ryan, T.P. (2000), *Statistical Methods for Quality Improvement*, 2nd ed., John Wiley and Sons, New York.

(24) Hayashi, M. et al. (2011), Compilation and use of genetic toxicity historical control data, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, Vol. 723/2, pp. 87-90.

(25) Rothfuss, A. et al. (2011), Improvement of in vivo genotoxicity assessment: combination of acute tests and integration into standard toxicity testing, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 723/2, pp. 108-120.

(26) Hayashi, M. et al. (1994), In vivo rodent erythrocyte micronucleus assay, *Mutation Research/Environmental Mutagenesis and Related Subjects*, Vol. 312/3, pp. 293-304.

(27) Fielder, R.J. et al. (1992), Report of British Toxicology Society/UK Environmental Mutagen Society Working Group. Dose setting in in vivo mutagenicity assays, *Mutagenesis*, Vol. 7/5, pp. 313-319.

(28) OECD (2000), "Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals Used in Safety Evaluation", OECD Environment, Health and Safety Publications (EHS), Series on Testing and Assessment, No. 19, OECD Publishing, Paris.

(29) LeBaron, M.J. et al. (2013), Influence of counting methodology on erythrocyte ratios in the mouse micronucleus test, *Environmental and Molecular Mutagenesis*, Vol. 54/3, pp. 222-228.

(30) Higashikuni, N., S. Sutou (1995), An optimal, generalized sampling time of 30 +/- 6 h after double dosing in the mouse peripheral blood micronucleus test, *Mutagenesis*, Vol. 10/4, pp. 313-319.

(31) Hayashi, M. et al. (2000), In vivo rodent erythrocyte micronucleus assay. II. Some aspects of protocol design including repeated treatments, integration with toxicity testing, and automated scoring, *Environmental and Molecular Mutagenesis*, Vol. 35/3, pp. 234-252.

(32) Asanami, S., K. Shimono (1997), High body temperature induces micronuclei in mouse bone marrow, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 390/1-2, pp. 79-83.

(33) Asanami, S., K. Shimono, S. Kaneda (1998), Transient hypothermia induces micronuclei in mice, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 413/1, pp. 7-14.

(34) Spencer, P.J. et al. (2007), Induction of micronuclei by phenol in the mouse bone marrow: I. Association with chemically induced hypothermia, *Toxicological Sciences*, Vol. 97/1, pp. 120-127.

(35) Hayashi, M., T. Sofuni, M. Jr. Ishidate (1983), An application of Acridine Orange fluorescent staining to the micronucleus test, *Mutation Research Letters*, Vol. 120/4, pp. 241-247.

(36) MacGregor, J.T., C.M. Wehr, R.G. Langlois (1983), A simple fluorescent staining procedure for micronuclei and RNA in erythrocytes using Hoechst 33258 and pyronin Y, *Mutation Research*, Vol. 120/4, pp. 269-275.

(37) Romagna, F., C.D. Staniforth (1989), The automated bone marrow micronucleus test, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, Vol. 213/1, pp. 91-104.

(38) Sun, J.T., M.J. Armstrong, S.M. Galloway (1999), Rapid method for improving slide quality in the bone marrow micronucleus assay; an adapted cellulose column procedure, *Mutation Research*, Vol. 439/1, pp. 121-126.

(39) Miller, B.M., I.D. Adler (1990), Application of antikinetochore antibody staining (CREST staining) to micronuclei in erythrocytes induced in vivo, *Mutagenesis*, Vol. 5/4, pp. 411-415.

(40) Miller, B.M. et al. (1991), Classification of micronuclei in murine erythrocytes: immunofluorescent staining using CREST antibodies compared to in situ hybridization with biotinylated gamma satellite DNA, *Mutagenesis*, Vol. 6/4, pp. 297-302.

(41) Russo, A. (2002), PRINS tandem labeling of satellite DNA in the study of chromosome damage, *American Journal of Medical Genetics*, Vol. 107/2, pp. 99-104.

(42) Gollapudi, B.B., L.G. McFadden (1995), Sample size for the estimation of polychromatic to normochromatic erythrocyte ratio in the bone marrow micronucleus test, *Mutation Research*, Vol. 347/2, pp. 97-99.

(43) OECD (2014), "Statistical analysis supporting the revision of the genotoxicity Test Guidelines", OECD Environment, Health and Safety Publications (EHS), Series on Testing and Assessment, No. 198, OECD Publishing, Paris.

(44) Richold, M. et al. (1990), "*In Vivo* Cytogenetics Assays", in *Basic Mutagenicity Tests, UKEMS Recommended Procedures. UKEMS Subcommittee on Guidelines for Mutagenicity Testing. Report. Part I revised*, Kirkland, D.J. (ed.), Cambridge University Press, Cambridge, pp. 115-141.

(45) Lovell, D.P. et al. (1989), "Statistical Analysis of *In Vivo* Cytogenetic Assays", in *Statistical Evaluation of Mutagenicity Test Data. UKEMS SubCommittee on Guidelines for Mutagenicity Testing, Report, Part III*, Kirkland, D.J. (ed.), Cambridge University Press, Cambridge, pp. 184-232.

(46) Hayashi, M. et al. (1994), Statistical analysis of data in mutagenicity assays: rodent micronucleus assay, *Environmental Health Perspectives*, Vol. 102/Suppl 1, pp. 49-52.

(47) Kim, B.S., M. Cho, H.J. Kim (2000), Statistical analysis of in vivo rodent micronucleus assay, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 469/2, pp. 233-241.

ANNEX 1

DEFINITIONS

Centromere: Region(s) of a chromosome with which spindle fibers are associated during cell division, allowing orderly movement of daughter chromosomes to the poles of the daughter cells.

Erythroblast: An early stage of erythrocyte development, immediately preceding the immature erythrocyte, where the cell still contains a nucleus.

Kinetochore: The protein structure that forms on the centromere of eukaryotic cells, which links the chromosome to microtubule polymers from the mitotic spindle during mitosis and meiosis and functions during cell division to pull sister chromatids apart.

Micronuclei: Small nuclei, separate from and additional to the main nuclei of cells, produced during telophase of mitosis (meiosis) by lagging chromosome fragments or whole chromosomes.

Normochromatic or mature erythrocyte: A fully matured erythrocyte that has lost the residual RNA that remains after enucleation and/or has lost other short-lived cell markers that characteristically disappear after enucleation following the final erythroblast division.

Polychromatic or immature erythrocyte: A newly formed erythrocyte in an intermediate stage of development, that stains with both the blue and red components of classical blood stains such as Wright's Giemsa because of the presence of residual RNA in the newly-formed cell. Such newly formed cells are approximately the same as reticulocytes, which are visualised using a vital stain that causes the residual RNA to clump into a reticulum. Other methods, including monochromatic staining of RNA with fluorescent dyes or labeling of short-lived surface markers such as CD71 with fluorescent antibodies, are now often used to identify the newly formed red blood cell. Polychromatic erythrocytes, reticulocytes, and CD71-positive erythrocytes are all immature erythrocytes, though each has a somewhat different age distribution.

Reticulocyte: A newly formed erythrocyte stained with a vital stain that causes residual cellular RNA to clump into a characteristic reticulum. Reticulocytes and polychromatic erythrocytes have a similar cellular age distribution.

ANNEX 2

THE FACTORIAL DESIGN FOR IDENTIFYING SEX DIFFERENCES IN THE *IN VIVO* MICRONUCLEUS ASSAY

The factorial design and its analysis

In this design, a minimum of 5 males and 5 females are tested at each concentration level resulting in a design using a minimum of 40 animals (20 males and 20 females, plus relevant positive controls).

The design, which is one of the simpler factorial designs, is equivalent to a two-way analysis of variance with sex and concentration level as the main effects. The data can be analysed using many standard statistical software packages such as SPSS, SAS, STATA, Genstat as well as using R.

The analysis partitions the variability in the dataset into that between the sexes, that between the concentrations and that related to the interaction between the sexes and the concentrations. Each of the terms is tested against an estimate of the variability between the replicate animals within the groups of animals of the same sex given the same concentration. Full details of the underlying methodology are available in many standard statistical textbooks (see references) and in the 'help' facilities provided with statistical packages.

The analysis proceeds by inspecting the sex x concentration interaction term in the ANOVA table[1]. In the absence of a significant interaction term the combined values across sexes or across concentration levels provide valid statistical tests between the levels based upon the pooled within group variability term of the ANOVA.

The analysis continues by partitioning the estimate of the between concentrations variability into contrasts which provide for a test for linear and quadratic contrasts of the responses across the concentration levels. When there is a significant sex x concentration interaction this term can also be partitioned into linear x sex and quadratic x sex interaction contrasts. These terms provide tests of whether the concentration responses are parallel for the two sexes or whether there is a differential response between the two sexes.

The estimate of the pooled within group variability can be used to provide pair-wise tests of the difference between means. These comparisons could be made between the means for the two sexes and between the means for the different concentration level such as for comparisons with the negative control levels. In those cases where there is a significant interaction comparisons can be made between the means of different concentrations within a sex or between the means of the sexes at the same concentration.

References

There are many statistical textbooks which discuss the theory, design, methodology, analysis and interpretation of factorial designs ranging from the simplest two factor analyses to the more complex forms

[1] Statisticians who take a modelling approach such as using General Linear Models (GLMs) may approach the analysis in a different but comparable way but will not necessarily derive the traditional anova table, which dates back to algorithmic approaches to calculating the statistics developed in a pre-computer age.

used in Design of Experiment methodology. The following is a non-exhaustive list. Some books provide worked examples of comparable designs, in some cases with code for running the analyses using various software packages.

Box, G.E.P, Hunter, W.G. and Hunter, J.S. (1978). Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building. New York: John Wiley & Sons.

Box G.E.P. & Draper, N.R. (1987) Empirical model-building and response surfaces. John Wiley & Sons Inc.

Doncaster, C.P. & Davey, A.J.H. (2007) Analysis of Variance and Covariance: How to choose and Construct Models for the Life Sciences. Cambridge University Press.

Mead, R. (1990) The Design of Experiments. Statistical principles for practical application. Cambridge University Press.

Montgomery D.C. (1997) Design and Analysis of Experiments. John Wiley & Sons Inc.

Winer, B.J. (1971) Statistical Principles in Experimental Design. McGraw Hill.

Wu, C.F.J & Hamada, M.S. (2009) Experiments: Planning, Analysis and Optimization. John Wiley & Sons Inc.