OECD GUIDELINE FOR THE TESTING OF CHEMICALS
PROPOSAL FOR UPDATING TEST GUIDELINE 473

In Vitro Mammalian Chromosomal Aberration 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 473 was adopted in 1983. In 1997 a revised version was issued, based on scientific progress made to that date. Modifications within the current version reflect nearly thirty years of experience with this test. This Test Guideline is part of a series of Test Guidelines on genetic toxicology. A document presented as an Introduction to the Test Guidelines on genetic toxicology (34) can also be referred to and provides succinct and useful guidance to users of these Test Guidelines.

2. The purpose of the in vitro chromosomal aberration test is to identify agents that cause structural chromosomal aberrations in cultured mammalian cells (1)(2)(3). Structural aberrations may be of two types, chromosome or chromatid. Polyploidy (including endoreduplication) is a common finding in chromosome aberration assays in vitro (see Introduction document (34)). While aneugens can induce polyploidy, polyploidy alone does not indicate aneugenic potential and can simply indicate cell cycle perturbation or cytotoxicity (51). This test is not designed to measure aneuploidy. An in vitro micronucleus test (35) would be recommended for the detection of aneuploidy.

3. The in vitro chromosomal aberration test may employ cultures of cell lines or primary cell cultures of human or rodent origin. The cells used are selected on the basis of growth ability in culture, stability of the karyotype (including chromosome number), chromosome diversity and spontaneous frequency of chromosomal aberrations (39). At the present time, the available data do not allow firm recommendations to be made but suggest it is important, when evaluating chemical hazards to consider the p53 status, genetic (karyotype) stability, DNA repair capacity and origin (rodent versus human) of the cells chosen for testing (34). The users of this Test Guideline are thus encouraged to consider the influence of these and other cell characteristics on the performance of a cell line in detecting the induction of chromosomal aberrations, as knowledge evolves in this area.

4. Definitions used are provided in Annex 1.

INITIAL CONSIDERATIONS AND LIMITATIONS

5. Tests conducted in vitro generally require the use of an exogenous source of metabolic activation unless the cells are metabolically competent with respect to the substances being tested (34). The exogenous metabolic activation system does not entirely mimic in vivo conditions. Care should also be taken to avoid conditions that could lead to artifactual positive results, i.e. chromosome damage not caused by direct interaction between the test chemical and chromosomes such conditions include changes in pH or osmolality (6) (7) (8), interaction with the medium components (5) (49) or excessive levels of cytotoxicity (24) (52) (53) (54).

6. This test is used to detect chromosomal aberrations that may result from clastogenic events. The analysis of chromosomal aberration induction should be done on cells in metaphase. It is thus essential that cells
should reach mitosis both in treated and in untreated cultures. For manufactured nanomaterials, specific adaptations of this Test Guideline are needed (see MN GD, Introduction document (34)).

PRINCIPLE OF THE TEST

7. Cell cultures of human or other mammalian origin are exposed to the test chemical both with and without an exogenous source of metabolic activation unless cells with an adequate metabolizing capability are used (see paragraph 12 and Introduction document (34) for more information). At predetermined intervals after the start of exposure of cell cultures to the test chemical, they are treated with a metaphase-arresting substance (e.g. Colcemid® or colchicine), harvested, stained and metaphase cells are analysed microscopically for the presence of chromatid-type and chromosome-type aberrations.

DESCRIPTION OF THE METHOD

Preparations

Cells

8. A variety of cell lines (e.g. Chinese Hamster Ovary (CHO), Chinese Hamster lung V79, Chinese Hamster Lung (CHL)/IU, TK6) or primary cell cultures, including human cells or other mammalian peripheral blood lymphocytes, can be used (39). When primary cells are used, for animal welfare reasons, the use of primary cells from human origin should be considered where feasible and sampled in accordance with the human ethical principles and regulations. Human peripheral blood lymphocytes should be obtained from young (approximately 18-35 years of age), non-smoking individuals with no known illness or recent exposures to genotoxic agents (e.g. chemicals, ionizing radiations) at levels that would increase the background incidence of chromosomal aberrations. This would ensure the background incidence of chromosomal aberrations to be low and consistent. The baseline incidence of chromosomal aberrations increases with age and this trend is more marked in females than in males (9) (10). If cells from more than one donor are pooled for use, the number of donors should be specified. It is necessary to demonstrate that the cells have divided from the beginning of treatment with the test chemical to cell sampling. Cell cultures are maintained in an exponential cell growth phase (cell lines) or stimulated to divide (primary cultures of lymphocytes), to expose the cells at different stages of the cell cycle, since the sensitivity of cell stages to the test chemical may not be known. The primary cells that need to be stimulated with mitogenic agents in order to divide are generally no longer synchronized during exposure to the test chemicals (e.g. human lymphocytes after a 48-hour mitogenic stimulation). The use of synchronized cells during treatment is not recommended, but can be acceptable if justified.

Media and culture conditions

9. Appropriate culture medium and incubation conditions (culture vessels, humidified atmosphere of 5% CO₂ if appropriate, incubation temperature of 37°C) should be used for maintaining cultures. Cell lines should be checked routinely for the stability of the modal chromosome number and the absence of Mycoplasma contamination (39) (40), and cells should not be used if contaminated or if the modal chromosome number has changed. The normal cell cycle time of cell lines or primary cultures used in the testing laboratory should be established and should be consistent with the published cell characteristics (41).
Preparation of cultures

10. Cell lines: cells are propagated from stock cultures, seeded in culture medium at a density such that the cells in suspensions or in monolayers will continue to grow exponentially until harvest time (e.g. confluency should be avoided for cells growing in monolayers).

11. Lymphocytes: whole blood treated with an anti-coagulant (e.g. heparin) or separated lymphocytes are cultured (e.g. for 48 hours for human lymphocytes) in the presence of a mitogen (e.g. phytohaemagglutinin (PHA) for human lymphocytes) in order to induce cell division prior to exposure to the test chemical.

Metabolic activation

12. Exogenous metabolising systems should be used when employing cells which have inadequate endogenous metabolic capacity. The most commonly used system that is recommended by default, unless otherwise justified, is a co-factor-supplemented post-mitochondrial fraction (S9) prepared from the livers of rodents (generally rats) treated with enzyme-inducing agents such as Aroclor 1254 (11) (12) (13) or a combination of phenobarbital and β-naphthoflavone (14) (15) (16) (17) (18) (19). The latter combination does not conflict with the Stockholm Convention on Persistent Organic Pollutants (20) and has been shown to be as effective as Aroclor 1254 for inducing mixed-function oxidases (14) (15) (16) (18). The S9 fraction typically is used at concentrations ranging from 1 to 2% (v/v) but may be increased to 10% (v/v) in the final test medium. The use of products that reduce the mitotic index, especially calcium complexing products (42) should be avoided during treatment. The choice of type and concentration of exogenous metabolic activation system or metabolic inducer employed may be influenced by the class of chemical being tested. For a more detailed discussion on the use of metabolising systems, see the Introduction document (34).

Test chemical preparation

13. Solid test chemicals should be prepared in appropriate solvents and diluted, if appropriate, prior to treatment of the cells (see paragraph 22). Liquid test chemicals may be added directly to the test system and/or diluted prior to treatment of the test system. Gaseous or volatile chemicals should be tested by appropriate modifications to the standard protocols, such as treatment in sealed culture vessels (21) (22) (23). Preparations of the test chemical should be made just prior to treatment unless stability data demonstrate the acceptability of storage.

Test conditions

Solvents

14. The solvent should be chosen to optimize the solubility of the test chemical without adversely impacting the conduct of the assay, e.g. changing cell growth, affecting the integrity of the test material, reacting with culture vessels, impairing the metabolic activation system. It is recommended that, wherever possible, the use of an aqueous solvent (or culture medium) should be considered first. Well established solvents are for example water or dimethyl sulfoxide. Generally organic solvents should not exceed 1% (v/v) and aqueous solvents (saline or water) should not exceed 10% (v/v) in the final treatment medium. If other than well-established solvents are used (e.g. ethanol or acetone), their use should be supported by
data indicating their compatibility with the test chemical, the test system and their lack of genetic toxicity at the concentration used. In the absence of that supporting data, it is important to include untreated controls (see Annex 1) to demonstrate that no deleterious or clastogenic effects are induced by the chosen solvent.

**Measuring cell proliferation and cytotoxicity and choosing treatment concentrations**

15. When determining the highest test chemical concentration, concentrations that have the capability of producing artifactual positive responses, such as those producing excessive cytotoxicity (see paragraph 21), precipitation in the culture medium (see paragraph 22), or marked changes in pH or osmolality (see paragraph 5), should be avoided. If the test chemical causes a marked change in the pH of the medium at the time of addition, the pH might be adjusted by buffering the final treatment medium so as to avoid artifactual positive results and to maintain appropriate culture conditions.

16. Measurements of cell proliferation are made to assure that sufficient treated cells have reached mitosis during the test and that the treatments are conducted at appropriate levels of cytotoxicity (see paragraphs 17 and 21). Cytotoxicity should be determined with and without metabolic activation in the main experiment using an appropriate indication of cell death and growth. While the evaluation of cytotoxicity in an initial test may be useful to better define the concentrations to be used in the main experiment, an initial test is not mandatory. If performed, it should not replace the measurement of cytotoxicity in the main experiment.

17. Relative Population Doubling (RPD) or Relative Increase in Cell Count (RICC) are appropriate methods for the assessment of cytotoxicity in cytogenetic tests (24) (25) (26) (see Annex 2 for formulas). In case of long-term treatment and sampling times after the end of treatment longer than 1.5 normal cell cycle length (i.e. longer than 3 cell cycle length in total), RPD might underestimate cytotoxicity (27). Under these circumstances RICC could be a better measure or the evaluation of cytotoxicity after 1.5 normal cell cycle length would be a helpful estimate using RPD.

18. For lymphocytes in primary cultures, while the mitotic index (MI) is a measure of cytotoxic/cytostatic effects, it is also influenced by the time after treatment, the mitogen used and possible cell cycle disruption. However, the MI is acceptable because other cytotoxicity measurements may be cumbersome and impractical.

19. While RICC and RPD for cell lines and MI for primary culture of lymphocytes are the recommended cytotoxicity parameters, other indicators (e.g. cell integrity, apoptosis, necrosis) could provide useful additional information.

20. At least three test concentrations (not including the solvent and positive controls) that meet the acceptability criteria (appropriate cytotoxicity, number of cells, etc) should be evaluated. Either replicate or single treated cultures may be used at each concentration tested. The results obtained in the independent replicate cultures at a given concentration can be pooled for the data analysis (28). The total number of cells analysed per concentration should remain the same either using replicates or single cultures, to provide confidence in the evaluation (see paragraph 30). For chemicals demonstrating little or no cytotoxicity, concentration intervals of approximately 2 to 3 fold will usually be appropriate. Where cytotoxicity occurs, the test concentrations selected should cover a range from that producing cytotoxicity as described in paragraph 21 and including concentrations at which there is moderate and little or no cytotoxicity. Many test chemicals exhibit steep concentration response curves and in order to obtain data at low and moderate cytotoxicity or to study the dose response relationship in detail, it will be necessary to
use more closely spaced concentrations and/or more than three concentrations (single cultures or replicates), in particular in situations where a repeat experiment is required (see paragraph 46).

21. If the maximum concentration is based on cytotoxicity, the highest concentration should aim to achieve $55 \pm 5\%$ cytotoxicity using the recommended cytotoxicity parameters (i.e. reduction in RICC and RPD for cell lines and reduction in MI for primary cultures of lymphocytes to $45 \pm 5\%$ of the concurrent negative control). Care should be taken in interpreting positive results only to be found in the higher end of this range (24).

22. For poorly soluble compounds that are not cytotoxic at concentrations lower than the lowest insoluble concentration, the highest concentration analysed should produce turbidity or a precipitate visible by eye or with the aid of an inverted microscope at the end of the treatment with the test chemical. Even if cytotoxicity occurs above the lowest insoluble concentration, it is advisable to test at only one concentration producing turbidity or with a visible precipitate because artifactual effects may result from the precipitate. At the concentration producing a precipitate, care should be taken to assure that the precipitate does not interfere with the conduct of the test (e.g. staining or scoring). The determination of solubility in the culture medium prior to the experiment may also be useful.

23. If no cytotoxicity or precipitate is observed, the highest test concentration should correspond to $10\,$mM, $2\,$mg/mL, or $2\,$µl/mL, whichever is the lowest (37) (43) (44). Where the test chemical is not of defined composition e.g. substance of unknown or variable composition, complex reaction products or biological materials (i.e. UV-VCBs) (38), environmental extracts etc., the top concentration of the mixture should be higher (e.g. at least $5\,$mg/ml) to increase the concentration of each of the components. It should be noted however that these requirements may differ for human pharmaceuticals (36).

**Controls**

24. Concurrent negative controls (see paragraph 14), consisting of solvent alone in the treatment medium and treated in the same way as the treatment cultures, should be included for every harvest time.

25. Positive controls are needed to demonstrate the ability of the cells used to identify clastogens under the conditions of the test protocol used and the proficiency of the exogenous metabolic activation system, when applicable (examples of positive controls are given in the table 1 below). Because *in vitro* mammalian cell tests for genetic toxicity are sufficiently standardized, the use of positive controls may be confined to a clastogen requiring metabolic activation. Provided it is done concurrently with the non-activated test using the same treatment duration, this single positive control response will demonstrate both the activity of the metabolic activation system and the responsiveness of the test system. Long term treatment (without S9) should however have its own positive control as the treatment duration will differ from the activated test. Each positive control should be used at one or more concentrations expected to give reproducible and detectable increases over background in order to demonstrate the sensitivity of the test system (i.e. the effects are clear but do not immediately reveal the identity of the coded slides to the reader), and the response should not be compromised by cytotoxicity exceeding the limits specified in the TG.
Table 1. Reference chemicals recommended for assessing laboratory performance and for selection of positive controls

<table>
<thead>
<tr>
<th>Category</th>
<th>Chemical</th>
<th>CASRN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clastogens active without metabolic activation</td>
<td>Methyl methanesulphonate</td>
<td>66-27-3</td>
</tr>
<tr>
<td></td>
<td>Mitomycin C</td>
<td>50-07-7</td>
</tr>
<tr>
<td></td>
<td>4-Nitroquinoline-N-Oxide</td>
<td>56-57-5</td>
</tr>
<tr>
<td></td>
<td>Cytosine arabinoside</td>
<td>147-94-4</td>
</tr>
<tr>
<td>2. Clastogens requiring metabolic activation</td>
<td>Benzo(a)pyrene</td>
<td>50-32-8</td>
</tr>
<tr>
<td></td>
<td>Cyclophosphamide</td>
<td>50-18-0</td>
</tr>
</tbody>
</table>

PROCEDURE

Treatment with test chemical

26. Proliferating cells are treated with the test chemical in the presence and absence of a metabolic activation system.

Culture harvest time

27. For thorough evaluation, which would be needed to conclude a negative outcome, all three following experimental conditions should be conducted using a short term treatment with and without metabolic activation and long term treatment without metabolic activation (see paragraphs 42, 43 and 44):

- Cells should be exposed to the test chemical without metabolic activation for 3-6 hours, and sampled at a time equivalent to about 1.5 normal cell cycle length after the beginning of treatment (18),

- Cells should be exposed to the test chemical with metabolic activation for 3-6 hours, and sampled at a time equivalent to about 1.5 normal cell cycle length after the beginning of treatment (18),

- Cells should be continuously exposed without metabolic activation until sampling at a time equivalent to about 1.5 normal cell cycle length. Certain chemicals (e.g. nucleoside analogues) may be more readily detected by treatment/sampling times longer than 1.5 normal cell cycle length (24).

In the event that any of the above experimental conditions lead to a positive response, it may not be necessary to investigate the remaining conditions.

Chromosome preparation

28. Cell cultures are treated with Colcemid® or colchicine usually for one to three hours prior to harvesting. Each cell culture is harvested and processed separately for the preparation of chromosomes. Chromosome preparation involves hypotonic treatment of the cells, fixation and staining (1). For cells in monolayer, care
should be taken to avoid loss of mitotic cells during removal of culture medium and washing. In such cases, medium or washing fluid should be centrifuged and the cells added back to the culture.

**Analysis**

29. All slides, including those of positive and negative controls, should be independently coded before microscopic analysis for chromosomal aberrations. Since fixation procedures often result in a proportion of metaphase cells with loss of chromosomes, the cells scored should, therefore, contain a number of centromeres equal to the modal number +/- 2 for all cell types.

30. At least 300 well-spread metaphases should be scored per concentration and control to conclude a test chemical as clearly negative (see paragraph 40). The 300 cells should be equally divided among the replicates, if applicable. In case of single culture per concentration (see paragraph 20), at least 300 well spread metaphases should be scored in this single culture. Scoring 300 cells has the advantage of increasing the statistical power of the test and in addition, zero values will be rarely observed (only 5%). This number can be reduced when high numbers of cells with chromosome aberrations are observed and the test chemical considered as clearly positive.

31. Cells with structural chromosomal aberration(s) including and excluding gaps should be scored. Chromatid- and chromosome-type aberrations should be recorded separately and classified by sub-types (breaks, exchanges). Procedures in use in the laboratory should ensure that analysis of chromosomal aberrations is performed by well-trained scorers and peer-reviewed if appropriate.

32. Although the purpose of the test is to detect structural chromosomal aberrations, it is important to record polyploidy and endoreduplication frequencies when these events are seen (see paragraph 2 and Introduction document (34) for interpretation).

**Proficiency of the laboratory**

33. In order to establish sufficient experience with the test prior to using it for routine testing, the laboratory should have performed a series of experiments with reference positive chemicals acting via different mechanisms and various negative controls (using various solvents/vehicle).

34. A selection of positive control chemicals (see Table 1 in paragraph 25) should be investigated with short and long treatments in the absence of metabolic activation, and also with short treatment in the presence of metabolic activation, in order to demonstrate proficiency to detect clastogenic compounds and determine the effectiveness of the metabolic activation system. The concentrations of the selected chemicals should be chosen so as to give reproducible and concentration-related increases above the background in order to demonstrate the sensitivity and dynamic range of the test system.

**Historical control data**

35. The laboratory should establish:
- A historical positive control range and distribution,
- A historical negative (untreated, solvent) control range and distribution.

36. When first acquiring data for an 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 control distribution, concurrent negative controls should ideally be within the 95% control limits of that distribution. The laboratory’s historical negative control database should initially be built with a minimum of 10 experiments 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 (45)), to identify how variable their positive and negative control data are, and to show that the methodology is ‘under control’ in their laboratory (50). 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 (30).

37. Any changes to the experimental protocol should be considered in terms of their consistency with the laboratory’s existing historical control databases. Any major inconsistencies should result in the establishment of a new historical control database.

38. Negative control data should consist of the incidence of cells with chromosome aberrations from a single culture or the sum of replicate cultures as described in paragraph 20. 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 limit they may be acceptable for inclusion in the historical control distribution as long as these data are not extreme outliers and there is no evidence that the test system is no longer ‘under control’ (see paragraph 36) and no evidence of technical or human failure.

DATA AND REPORTING

Presentation of the results

39. The percentage of cells with structural chromosomal aberration(s) should be evaluated. Chromatid- and chromosome-type aberrations classified by sub-types (breaks, exchanges) should be listed separately with their numbers and frequencies for experimental and control cultures. Gaps are recorded and reported separately but not included in the total aberration frequency. Percentage of polyploidy and/or endoreduplicated cells are reported when seen.

40. Concurrent measures of cytotoxicity for all treated, negative and positive control cultures in the main aberration experiment(s) should be recorded.

41. Individual culture data should be provided. Additionally, all data should be summarised in tabular form.

Acceptability Criteria

42. Acceptance of a test is based on the following criteria:

- The concurrent negative control is considered acceptable for addition to the laboratory historical negative control database as described in paragraph 38.

- Concurrent positive controls (see paragraph 25) should induce responses that are compatible with those generated in the historical positive control data base and produce a statistically significant increase compared with the concurrent negative control.

- Cell proliferation criteria in the solvent control should be fulfilled (Paragraphs 16 and 17), e.g. a mitotic index of at least 5% or a PD of 1.4.
- All three experimental conditions were tested unless one resulted in positive results (see paragraph 27).
- Adequate number of cells and concentrations should be analysable (paragraphs 30 and 20).
- The criteria for the selection of top concentration are consistent with those described in paragraphs 21, 22 and 23.

**Evaluation and interpretation of results**

43. Providing that all acceptability criteria are fulfilled, a test chemical is considered to be clearly positive if, in any of the experimental conditions examined (see paragraph 27):
   a) at least one of the test concentrations exhibits a statistically significant increase compared with the concurrent negative control,
   b) the increase is dose-related when evaluated with an appropriate trend test,
   c) any of the results are outside the distribution of the historical negative control data (e.g. Poisson-based 95% control limit; see paragraph 39).

When all of these criteria are met, the test chemical is then considered able to induce chromosomal aberrations in cultured mammalian cells in this test system. Recommendations for the most appropriate statistical methods can be found in the literature (46) (47) (48).

44. Providing that all acceptability criteria are fulfilled, a test chemical is considered clearly negative if, in all experimental conditions examined (see paragraph 27):
   a) none of the test concentrations exhibits a statistically significant increase compared with the concurrent negative control,
   b) there is no concentration-related increase when evaluated with an appropriate trend test
   c) all results are inside the distribution of the historical negative control data (e.g. Poisson-based 95% control limit; see paragraph 39).

The test chemical is then considered unable to induce chromosomal aberrations in cultured mammalian cells in this test system.

45. There is no requirement for verification of a clearly positive or negative response.

46. In case the response is neither clearly negative nor clearly positive as described above or in order to assist in establishing the biological relevance of a result, the data should be evaluated by expert judgement and/or further investigations. Scoring additional cells (where appropriate) or performing a repeat experiment possibly using modified experimental conditions (e.g. concentration spacing, other metabolic activation conditions [i.e. S9 concentration or S9 origin]) could be useful.

47. In rare cases, even after further investigations, the data set will preclude making a conclusion of positive or negative results, and will therefore be concluded as equivocal.

48. An increase in the number of polyploid cells may indicate that the test chemical has the potential to inhibit mitotic processes and to induce numerical chromosomal aberrations (31). An increase in the number of cells with endoreduplicated chromosomes may indicate that the test chemical has the potential to inhibit cell cycle progress (32)(33), which is a different mechanism of inducing numerical chromosome changes.
than inhibition of mitotic processes (see paragraph 2). Therefore incidence of polyploid cells and cells with endoreduplicated chromosomes should be recorded separately.

Test report

49. The test report must include the following information:

Test chemical:
- identification data and CAS no., if known;
- physical nature and purity;
- physicochemical properties relevant to the conduct of the study;
- stability of the test chemical itself, if known;
- solubility and stability of the test chemical in solvent, if known.
- measurement of pH, osmolality and precipitate in the culture medium added with the test chemical, as appropriate.

Solvent:
- justification for choice of solvent.
- percentage of solvent in the final culture medium should also be indicated.

Cells:
- type and source of cells
- karyotype features and suitability of the cell type used;
- absence of mycoplasma, in case of cell lines;
- for cell lines, information on cell cycle length, doubling time or proliferation index;
- sex of blood donors, age and any relevant information on the donor, whole blood or separated lymphocytes, mitogen used;
- number of passages, if available, in case of cell lines;
- methods for maintenance of cell cultures, in case of cell lines;
- modal number of chromosomes, in case of cell lines.

Test conditions:
- identity of metaphase arresting substance, its concentration and duration of cell exposure;
- concentration of test chemical expressed as final concentration in the culture medium (e.g. µg or mg/mL, or mM of culture medium).
- rationale for selection of concentrations and number of cultures including, e.g. cytotoxicity data and solubility limitations;
- composition of media, CO₂ concentration if applicable, humidity level;
- concentration (and/or volume) of solvent and test chemical added in the culture medium;
- incubation temperature;
- incubation time;
- duration of treatment;
- harvest time after treatment;
- cell density at seeding, if appropriate;
- type and composition of metabolic activation system (source of S9, method of preparation of the S9 mix, the concentration or volume of S9 mix and S9 in the final culture medium, quality controls of S9);
- positive and negative control compounds, final concentrations for each conditions of treatment;
- methods of slide preparation and staining technique used;
- criteria for acceptability of assays;
- criteria for scoring aberrations;
- number of metaphases analyzed;
- methods for the measurements of cytotoxicity;
- any supplementary information relevant to cytotoxicity and method used;
- criteria for considering studies as positive, negative or equivocal;
- methods used to determine pH, osmolality and precipitation.

Results:

- the number of cells treated and the number of cells harvested for each culture in case of cell lines
- cytotoxicity measurements, e.g. RPD, RICC, MI, other observations if any;
- information on cell cycle length, doubling time or proliferation index in case of cell lines;
- signs of precipitation and time of the determination;
- definition for aberrations, including gaps;
- Number of cells scored, number of cells with chromosomal aberrations and type of chromosomal aberrations given separately for each treated and control culture, including and excluding gaps;
- changes in ploidy (polyploid cells and cells with endoreduplicated chromosomes, given separately) if seen;
- concentration-response relationship, where possible;
- concurrent negative (solvent) and positive control data (concentrations and solvents);
- historical negative (solvent) and positive control data, with ranges, means and standard deviations and confidence interval (e.g. 95%) as well as the number of data;
- statistical analyses, p-values if any.

Discussion of the results.

Conclusions.
LITERATURE


(26) Sheila Galloway, Elisabeth Lorge, Marilyn J. Aardema, David Eastmond, Mick Fellows, Robert Heflich, David Kirkland, Dan D. Levy, Anthony M. Lynch, Daniel Marzin, Takeshi Morita, Maik


(34) OECD (2013), Draft Introduction to the OECD guidelines on genetic toxicology testing and guidance on the selection and application of assays


(36) International Conference on Harmonisation (ICH) Guidance S2 (R1) on Genotoxicity Testing and Data Interpretation For Pharmaceuticals Intended For Human Use

(37) OECD (2013) Document supporting the WNT decision to implement revised criteria for the selection of the top concentration in the *in vitro* mammalian cell assays on genotoxicity (Test Guidelines 473, 476 and 487) Olis cote:to be added


(39) ILSI paper (draft), Lorge, E., M. Moore, J. Clements, M. O Donovan, F. Darroudi, M. Honma, A. Czich, J van Benthem, S. Galloway, V. Thybaud, B. Gollapudi, M. Aardema, J. Kim, D.J. Kirkland, Recommendations for good cell culture practices in genotoxicity testing.

(40) Sandra Coecke, Michael Balls, Gerard Bowe, John Davis, Gerhard Gstraunthaler, Thomas Hartung, Robert Hay, Otto-Wilhelm Merten, Anna Price, Leonard Schechtman, Glyn Stacey and


(50) OECD (2013) Draft report on statistical issues related to OECD in vitro genotoxicity Test Guidelines


(52) D. Kirkland, M. Aardema, L. Henderson, L. Müller, Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. I: Sensitivity, specificity and relative predictivity, Mutat. Res. 584 (2005) 1–256.

Annex 1

DEFINITIONS

Aneuploidy: any deviation from the normal diploid (or haploid) number of chromosomes by a single chromosome or more than one, but not by entire set(s) of chromosomes (polyploidy).

Apoptosis: programmed cell death characterized by a series of steps leading to a disintegration of cells into membrane-bound particles that are then eliminated by phagocytosis or by shedding.

Cell proliferation: increase in cell number as a result of mitotic cell division.

Chromatid break: discontinuity of a single chromatid in which there is a clear misalignment of one of the chromatids.

Chromatid gap: non-staining region (achromatic lesion) of a single chromatid in which there is minimal misalignment of the chromatid.

Chromatid-type aberration: structural chromosome damage expressed as breakage of single chromatids or breakage and reunion between chromatids.

Chromosome-type aberration: structural chromosome damage expressed as breakage, or breakage and reunion, of both chromatids at an identical site.

Chromosome diversity: diversity of chromosome shapes (e.g. metacentric, acrocentric, etc..) and sizes.

Clastogen: any substance which causes structural chromosomal aberrations in populations of cells or organisms.

Concentrations: refer to final concentrations in the culture medium.

Cytotoxicity: For the assays covered in this guideline using cell lines, cytotoxicity is identified as a reduction in relative population doubling (RPD) or relative increase in cell count (RICC) of the treated cells as compared to the negative control (see paragraph 17 and Annex 2). For the assays covered in this guideline using primary cultures of lymphocytes, cytotoxicity is identified as a reduction in mitotic index (MI) of the treated cells as compared to the negative control (see paragraph 18 and Annex 2).

Endoreduplication: a process in which after an S period of DNA replication, the nucleus does not go into mitosis but starts another S period. The result is chromosomes with 4, 8, 16, ...chromatids.

Genotoxic: a general term encompassing all types of DNA or chromosome damage, including breaks, adducts, rearrangements, mutations, chromosome aberrations, and aneuploidy. Not all types of genotoxic effects result in mutations or stable chromosome damage.

Mitotic index (MI): the ratio of cells in metaphase divided by the total number of cells observed in a population of cells; an indication of the degree of proliferation of that population.
Mitosis: division of the cell nucleus usually divided into prophase, prometaphase, metaphase, anaphase and telophase.

Mutagenic: produces a heritable change of DNA base-pair sequences(s) in genes or of the structure of chromosomes (chromosome aberrations).

Numerical aberration: a change in the number of chromosomes from the normal number characteristic of the cells utilised.

Polyploidy: numerical chromosomal aberrations in cells or organisms involving entire set(s) of chromosomes, as opposed to an individual chromosome or chromosomes (aneuploidy).

p53 status: p53 protein is involved in cell cycle regulation, apoptosis and DNA repair. Cells deficient in functional p53 protein, unable to arrest cell cycle or to eliminate damaged cells via apoptosis or other mechanisms (e.g. induction of DNA repair) related to p53 functions in response to DNA damage induction, should be theoretically more prone to gene mutations or chromosomal aberrations.

Relative Increase in Cell Counts (RICC): the increase in the number of cells in chemically-exposed cultures versus increase in non-treated cultures, a ratio expressed as a percentage.

Relative Population Doubling (RPD): the increase in the number of population doublings in chemically-exposed cultures versus increase in non-treated cultures, a ratio expressed as a percentage.

S9 liver fraction: supernatant of liver homogenate after 9000g centrifugation, i.e., raw liver extract.

S9 mix: mix of the S9 liver fraction and cofactors necessary for metabolic activation added in buffer.

Solvent control: General term to define the control cultures receiving the solvent alone used to dissolve the test chemical.

Structural aberration: a change in chromosome structure detectable by microscopic examination of the metaphase stage of cell division, observed as deletions and fragments, intrachanges or interchanges.

Untreated controls: cultures that receive no treatment (ie no test chemical nor solvent) but are processed the same way as the cultures receiving the test chemical.
ANNEX 2

FORMULAS FOR CYTOTOXICITY ASSESSMENT

Mitotic index (MI):

\[
\text{MI} \, (\%) = \frac{\text{Number of mitotic cells}}{\text{Total number of cells scored}} \times 100
\]

Relative Increase in Cell Counts (RICC) or Relative Population Doubling (RPD) is recommended, as both take into account the proportion of the cell population which has divided.

\[
\begin{align*}
\text{RICC} \, (\%) &= \frac{(\text{Increase in number of cells in treated cultures (final} - \text{starting})}{(\text{Increase in number of cells in control cultures (final} - \text{starting})}) \times 100 \\
\text{RPD} \, (\%) &= \frac{(\text{No. of Population doublings in treated cultures})}{(\text{No. of Population doublings in control cultures})} \times 100
\end{align*}
\]

where:

Population Doubling = \[\log (\text{Post-treatment cell number ÷ Initial cell number})] ÷ \log 2

As an example, a RICC, or a RPD of 53% indicates 47% cytotoxicity/cytostasis.

In any case, the number of cells before treatment should be measured and the same for treated and negative control cultures.

For information relative cell counts (RCC, i.e. Number of cells in treated cultures/ Number of cells in control cultures) used as cytotoxicity parameter in the past is no more recommended, because it is considered to possibly underestimate the cytotoxicity.
**Annex** (to support the recommendations in the TG – working document to be deleted at the final stage)

**STATISTICAL CONSIDERATIONS ON THE NUMBER OF CELLS TO SCORE**

A statistical analysis was performed to derive recommendations on these statistical issues (50).

Scoring 300 cells per concentration was considered a compromise between achieving an optimal number for the statistical power of the assay and the practicability of scoring. It also takes into account that where the events (cells with a chromosomal aberration) have an incidence of 1% (p=0.01) and are Poisson distributed and 500 cells (n=500) are scored per cultures then cultures with no cells with aberrations will be rare. This is based upon the Poisson distribution of counts with a mean (np) of 5 where less that 0.7% of cultures are expected to be zero. The Poisson distribution also approximates to a normal distribution when np ≥5. In the case of 300 cells, assuming a Poisson distribution approximately 5% of cultures would be expected to be zero.

**STATISTICAL CONSIDERATIONS ON THE USE OF HISTORICAL DATA**

A statistical analysis was performed to derive recommendations on these statistical issues (50).

Historical negative control data are used both for acceptability of the assays and to help with the interpretation of the results in those case where results are neither clearly negative nor clearly positive. Quality control methods (for example C-charts or X-bar charts) should be used to characterise the distribution of the negative control data base.

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 limit they may be acceptable for inclusion in the historical control distribution as long as these data are not extreme outliers and there is no evidence that the test system is no longer ‘under control’ (see paragraph 36) and no evidence of technical or human failure.

These methods can be used to assess the acceptability of the assay, using the 95% control limits. In order to avoid narrowing the database, all data should be included in the historical data base, even when rejected for the evaluation of the results, unless a technical or human failure has been identified. In this case the data will not be added to the database. The choice of the 95% control limit over the more stringent 99% control limit provides flexibility to avoid unnecessarily discarding experiments and avoids narrowing the control database range.

In the interpretation of the results, the treated values should be compared with the 95% control levels for the distribution of the historical negative control data base. Results which fall within the 95% control levels for the historical control could be considered broadly consistent with historical negative control data.