REPORT ON

MICRONUCLEUS TEST (MNT) *in vitro*

Validation Management Team:
Silvio Albertini[1], Jan van Benthem[2], Raffaella Corvi[3], Sebastian Hoffmann[3], Daniela Maurici[3], Stefan Pfuhler[4], Philippe Vanparys[5].

[1] Non-Clinical Drug Safety, F.Hoffmann-La Roche Ltd, Basel, Switzerland
[2] National Institute of Public Health and the Environment (RIVM), Bilthoven, NL
[3] European Centre for Validation of Alternative Methods (ECVAM), IHCP, European Commission JRC, Ispra (VA), Italy
[4] Product Safety and Regulatory Affairs, Wella Service GmbH, Darmstadt, Germany
[5] Mechanistic Toxicology, Toxicology and Pathology, Johnson & Johnson Pharmaceutical Research & Development, Beerse, Belgium

---

SUMMARY

In the past decade several studies comparing the *in vitro* chromosome aberration test (CAT) and the micronucleus test (MNT) *in vitro* were performed. A high correlation was observed (>85%). No formal ECVAM validation for the micronucleus *in vitro* assay has been performed. Therefore a working group was established to pool together the existing data, which would support the validity of the MNT *in vitro* (compared to the chromosome aberration *in vitro* assay) on the basis of a new proposed modular validation approach.

The primary focus was on the evaluation of the potential of the MNT *in vitro* as alternative to the standard chromosome aberration assay *in vitro*. The potential as alternative to the *in vivo* micronucleus assay was not assessed in-depth by the group. However, there is supportive data/evidence, that the MNT *in vitro* has a high predictivity for the *in vivo* MNT assay. This should be evaluated in a second phase.

The working group evaluated in a first step the available published data and came to the conclusion, that two publications met the criteria for a retrospective validation according to the criteria previously laid out by the group. The two publications were:

1. von der Hude W. et al. (2000) *In vitro* micronucleus assay with Chinese hamster V79 cells – results of a collaborative study with in situ exposure to 26 chemical substances. Mutation Research 468 (2), 137-163
2. Lorge E. et al. (2004) SFTG international collaborative study on *in vitro* micronucleus test I. General conditions and overall conclusions of the study. Submitted Mutation Research.

These two studies were evaluated in depth (including the re-analysis of raw data) and provided the information required for modules 2, 3 and 4 for the assessing the reliability (reproducibility) of the test. For the assessment of the concordance between the MNT *in vitro* and the CAT *in vitro* (module 5), additional published data were taken into consideration.

Based on this retrospective validation, the Validation Management Team (VMT) concluded that the MNT *in vitro* is reliable and reproducible and can therefore be used as an alternative method to the CA assay *in vitro.*

TABLE OF CONTENTS

RATIONALE FOR THE PROPOSED TEST

Introduction

An expert meeting on Micronucleus Test (MNT) *in vitro* was held at ECVAM on 13th-14th April 2004. During that meeting the expert group (see below for participants) decided that, due to the consistent amount of data available and the high interest in using the MNT *in vitro* for regulatory purposes, it was necessary to start a retrospective validation of the test. The expert group strongly recommended to support the scientific validity of the test by compiling a dossier based on existing data. In order to evaluate whether the test meets all data requirements requested by the ECVAM principles on test validity, it was decided to follow the modular approach.

The modular validation approach (Hartung et al., 2004) is defined by 7 validity modules:

1. Test definition
2. Within-laboratory variability
3. Transferability
4. Between-laboratory variability
5. Predictive capacity
6. Applicability domain
7. (Minimum performance standards)

Modules 1-4 cover the reproducibility aspects of an assay, 5 the predictivity/concordance, 6 the applicability domain of the test and module 7 defines the requirements to accept additional data/assays for the same endpoint. Module 7 will not be considered in this evaluation of retrospective data.

The current report, which was prepared by the ECVAM Validation Management Team (VMT), presents the outcome of the retrospective validation study. Available literature data and validation efforts were taken into account.

Participants at the first ECVAM meeting, April 2004

Silvio Albertini[1], Raffaella Corvi[2], Robin Fielder[3], Peggy J. Guzzie[4], Thomas Hartung[2], Sebastian Hoffmann[2], David Kirkland[5], Micheline Kirsch-Volders[6], Elisabeth Lorge[7], Daniela Maurici[2], Lucia Migliore[8], Hannu Norppa[9], Laurie Scott[2], Philippe Vanparys[10].

1 Non-Clinical Drug Safety, F.Hoffmann-La Roche Ltd, Basel, Switzerland
2 European Centre for Validation of Alternative Methods (ECVAM), IHCP, European Commission JRC, Ispra (VA), Italy
3 Department of Health, Environmental Chemical Unit PH5D, London SE1 6LH, UK
4 Safety Sciences, Pfizer, Groton, CT 06340, USA
5 Covance Laboratories Limited, Harrogate, North Yorkshire, England
6 Vrije Universiteit Brussel, Faculteit Wetenschappen, Laboratorium voor Cellulaire Genetica, 1050 Brussel, Belgium
7 Servier Group, Drug Safety Assessment, 905, Orleans-Gidy, France
8 Universita' di Pisa, Dept. of Human and Environmental Sciences, Section of Genetics, Pisa, Italy
9 Laboratory of Molecular and Cellular Toxicology, Department of Industrial Hygiene and Toxicology, Finnish Institute of Occupational Health, Helsinki, Finland
10 Mechanistic Toxicology, Toxicology and Pathology, Johnson & Johnson Pharmaceutical Research & Development, Beerse, Belgium

Nominated Use
Once validated the micronucleus test (MNT) *in vitro* is meant to be an alternative or replacement test method for the chromosome aberration test (CAT) *in vitro* (OECD TG 473; EC Annex V B.10). In the long term, the goal is that the MNT *in vitro* replaces the MNT *in vivo*. However, this issue was not addressed in the present report.

The CAT has been widely used and recommended by regulatory authorities for the assessment of chromosome damage. Metaphase analysis is very tedious and time-consuming. The *in vivo* MNT in bone marrow of rodents has long been established and used (OECD TG 474; EC Annex V B.12), as rapid alternative to the much more labour-intensive evaluation of chromosome aberration *in vivo*.

Despite a large amount of data available, the *in vitro* MNT assay is not yet generally accepted by most regulatory authorities as an alternative system in a test battery. A draft Test Guideline on MNT *in vitro* (TG 487) has already been submitted to the OECD in 2004 (see Annex 1). However, the OECD and some of its member countries agreed not to further process with the acceptance of the Test Guideline, until the ECVAM retrospective validation would be finalized.

Current Use
The test is currently being used by academics, pharmaceutical and cosmetic industry, and CROs as replacement of the *in vitro* CAT for internal hazard assessment and prioritization. In some instances it is already accepted by national regulatory authorities. Moreover, the test is included in the EC Technical Guideline Document on Risk Assessment (TGD), 2003.

Strategy Integration
The test will be used as alternative or replacement to the CAT *in vitro*, and therefore, it will be integrated in the current strategy for testing of chemicals as alternative test to the CAT *in vitro*. Although the MNT *in vitro* will initially be used for assessment of chemicals, it has also a potential in other areas (e.g. agrochemicals, pharmaceuticals).

Patents
The test method has not been patented.

Publications
An extensive amount of published data is available to support the validation of the *in vitro* MNT using various cell lines or human lymphocytes. These include, among others, the international validation studies coordinated by the French branch of the European Environmental Mutagen Society (SFTG) (submitted) and the German collaborative study on 26 chemicals (von der Hude et al., 2000). These data were considered at the 3rd International Workshop on Genotoxicity Testing (IWGT, Plymouth, USA) in June 2002 resulting in a report of the *in vitro* micronucleus assay working group's conclusion as outlined at the 3rd IWGT meeting (Kirsch-Volders et al., 2003; see Annex 2). At that workshop international experts from Japan, Europe and USA reviewed current methodologies and data for the *in vitro* MNT and consensus was reached on all the key points related to the protocols to be used. There was general agreement that the method had now been adequately validated at that point. Furthermore, it was considered a valid alternative to the CAT *in vitro*, and that it has several advantages, especially the possibility to detect aneugens.

Additional validation efforts also include studies carried out by the pharmaceutical industry (Albertini et al., 1997; Miller et al., 1997; Miller et al., 1998; von der Hude et al., 2000; Garriot et

al., 2000), an inter-laboratory validation study carried out under the auspices of the Japanese Ministry of Labour (Matsushima et al., 1999) and a literature review carried out by a GUM Working Group (Miller et al., 1998). In addition, the large database presented in a recent review (Kirkland et al., 2005) contains, among other genotoxicity results, a substantial amount of data created with the MNT *in vitro*.

MODULE 1 - TEST DEFINITION

The following sections provide information about the scientific purpose of the test as well as the expert's opinion regarding which protocol requirements and validation aspects do need to be fulfilled to allow the inclusion of data in the retrospective analysis according to the ECVAM modular validation approach (Hartung et al., 2004).

Scientific basis for the proposed test method

The proposed test method is intended to detect:
(1) clastogens more efficiently and with less investment in time and training than the *in vitro* CAT;
(2) aneugens not currently detected in regulatory *in vitro* genotoxicity tests.

Description of the endpoint predicted and the mechanistic basis of the test

The *in vitro* micronucleus assay is a mutagenicity test system which enables the detection of the potential of a test substance to induce the formation of small membrane-locked DNA fragments, i.e. micronuclei, in the cytoplasm of interphase cells. These micronuclei may originate from acentric fragments (chromosome fragments lacking a centromere), centric fragments (chromosome fragments containing a centromere) or whole chromosomes which are unable to migrate with the rest of the chromosomes during the anaphase of cell division. The assay thus has the potential to detect the activity of both clastogenic and aneugenic chemicals (Kirsch-Volders et al., 1997; Parry and Sorrs, 1993). By using the cytokinesis-block methodology (addition of the actin polymerization inhibitor cytochalasin B during the mitosis), it can be identified whether or not a cell has undergone cell division after the cells have been treated with a test substance. A cell which has gone through one cell division in the presence of cytochalasin B appears as binucleated cell. To demonstrate cell proliferation by using cytochalasin B is a prerequisite for primary cells like human lymphocytes because those are non-actively dividing cells, while cell lines can be tested with or without cytochalasin B. By immunochemical labeling of kinetochores or hybridization with general or chromosome specific centromeric/telomeric probes, the mechanism of micronucleus induction can be studied (clastogenicity vs. aneuploidy) (Fenech and Morley, 1986; Eastmond and Tucker, 1989; Eastmond and Pinkel, 1990; Miller et al., 1991; Farooqi et al., 1993; Migliore et al., 1993; Norppa et al., 1993; Eastmond et al., 1994; de Stoppelaar et al., 1997; 1999).
Further useful information can be gained by including additional evaluation criteria (other endpoints as well as mechanistic considerations). Endpoints such as necrosis, apoptosis, mitotic arrest or delay (mitotic death), as well as mechanistic insights such as C-mitosis, chromosome displacement, gene amplification, chromosome breakage and loss, centrosome abnormalitiy, chromosome rearrangements, dicentrics/rings, amitosis, cytokinesis abnormality can be determined/evaluated (see Figure 1) and can provide additional valuable information.

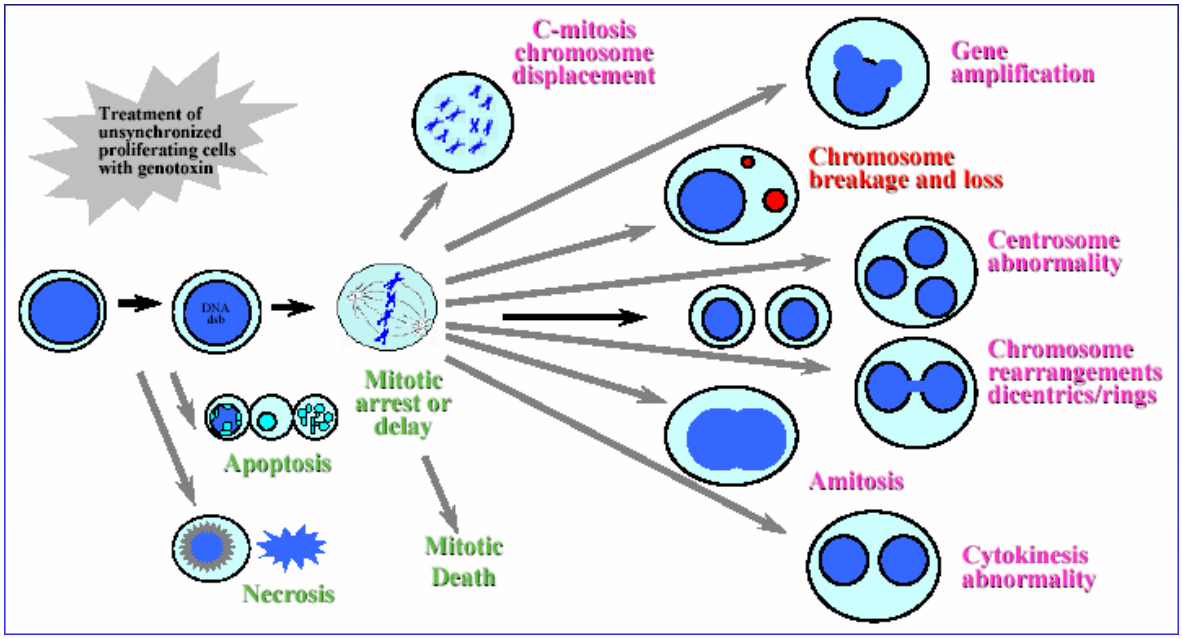

Figure 1: The MNT *in vitro*: A multi endpoint assay for co-detection of clastogenic and aneugenic activity, apoptosis/ necrosis, assessment of cell proliferation/cytostatic effects and mechanistic implications.

Advantages of MNT over CAT

The MNT has a number of advantages over metaphase analysis to measure chromosome damage.

1) Micronuclei in interphase cells can be assessed much more objectively than chromosomal aberrations in metaphase cells.

2) The training requirements for a person to be competent in scoring the slides are much less rigorous for MNT than for metaphase analysis.

3) As there is no requirement to count the chromosomes in a metaphase preparation, nor to evaluate subtle chromatid and chromosome damage, but only to determine whether or not a cell contains a micronucleus, the preparations can be scored much more quickly. This allows scoring thousands instead of hundreds of cells per treatment, which gives greater statistical power to the assay.

4) Since the micronuclei may contain whole (lagging) chromosomes, the MNT has the potential to detect aneuploidy-inducing agents which are currently very difficult to study in conventional CTA.

Definition of protocol requirements

The group defined criteria for all crucial aspects of a valid MNT *in vitro*. These criteria were defined as if they had to be applied, in a best case scenario, to a prospective validation study. Many aspects that were considered were taken from the well established information required for the CAT *in vitro* (OECD Test Guideline 473; EC Annex V B.10). Because the criteria were defined as rigorous as if they would have to apply to a prospective validation, it was unlikely that a single available study would fulfil all of them. In the end, only the studies that best fulfilled the most important acceptance criteria were taken into consideration for the retrospective evaluation of the

validity of the MNT *in vitro*. In addition, aspects to be considered for retrospective validation were included (e.g. number of compounds to be evaluated per class; inclusion of negative compounds).

- Cells

Any replicating mammalian cell, including those routinely used for clastogenicity testing, with a defined karyotype (diploid or near-diploid) and stable low background frequency of MN can be used (e.g. CHO, CHL, V79, L5178Y, human lymphocytes). Cells must be maintained and checked according to accepted good scientific practices (e.g. growth characteristics, lack of mycoplasma). Human lymphocytes should be taken from non-smoking, young, healthy donors.

- Media/culture conditions

Cells should be grown in appropriate media that support expected cell cycle times or plating efficiency. For inter-laboratory variability measures, media, cells and serum should be provided from a single source. To measure the robustness of the assay, participants may use the cells and media conditions routinely used within their facility; however treatment conditions (serum %, CO_2 etc.) will be specified.

- Preparation of cultures

Cell lines should be prepared from frozen stocks and allowed to achieve exponential growth at the time of treatment. They should be seeded at a density that will ensure they will not reach confluency before the time of harvest.

Treatment of human lymphocytes (whole blood or separated lymphocytes) should start not earlier than 24 hr after phytohemagglutinin (PHA) stimulation.

- Metabolic activation

S9 should be used from the liver of rodents induced with Aroclor 1254 or phenobarbitone/β-naphthoflavone, and the quality determined in terms of protein content and ability to activate known indirect genotoxins. Standard cofactors should be used. Final concentration of S9 in the treatment medium should range from 1-10% (v/v).

- Test substances

For effective evaluation of the scientific purpose of the *in vitro* MNT, a cumulative database of at least 10 clastogens and 5 aneugens, representing different mechanisms of action at the molecular level and different chemical classes, should be established. They should be selected according to their established mechanisms of action as determined by expert opinion/peer review. In addition, at least 5 substances that are toxic or non-toxic, but neither clastogenic nor aneugenic, will be included.

- Source

Chemicals may either be supplied from a central source or purchased from a recognized, reputable supplier. Sufficient information on purity and stability must be available. Chemicals should be stored in conditions (e.g. dry, dark, cool) according to manufacturer's specifications, and used within any specified expiry date. All chemicals should be handled as if they were carcinogenic.

- Preparation

Adequate instructions for preparation will be provided by the coordinator of the validation study. Solutions should be freshly prepared and used within the day of preparation, unless stability data

are provided which indicate that solutions should be used immediately, or can be aliquoted, frozen and used over a longer period.

- Coded/blinded

It is not necessary to code the compounds for treatment – in some cases it may be desirable for the technicians performing the treatments to know the physico-chemical properties of the chemical (volatility, immiscibility issues, etc.) – but it is mandatory that the slides are coded before analysis.

- Solvents/vehicles

Instructions will be provided on which solvent to use for each chemical, or whether to dissolve directly into culture medium. Substances dissolved in organic solvents (e.g. DMSO, ethanol, DMF, acetone) will be diluted in the treatment medium. Organic solvents will not exceed 1% (v/v) in the final treatment medium.

- Use of cytochalasin B

Use of cytochalasin B should be mandatory for studies with human lymphocytes, but it does not need to be used with cell lines as long as cell proliferation is assured (e.g. cell counting at start of treatment and time of harvest, or parallel cultures containing cytochalasin B are used) in all, including treated, cultures. The concentration of cytochalasin B should be in the range 3-6 µg/ml, and should achieve at least 50% binucleates in control cultures.

- Exposure concentration/toxicity

Toxicity can be indirectly assessed using a variety of methods. The following parameters are usually considered appropriate to assess cytotoxicity (by comparison with the respective control values):

- cytokinesis-block proliferation index (CBPI)
- replication or proliferation index
- % of binucleated cells
- % of multinucleated cells
- cell counting
- population doubling

Whereas a range of doses may be selected in a range-finding experiment, concurrent measures (if necessary in parallel cultures) of toxicity must be used to select the concentrations for analysis of MN. At least 3 analysable concentrations should be scored. Where toxicity is induced, the highest dose should induce substantial (at least 50%) toxicity, the middle dose should induce intermediate toxicity and the lowest dose should induce little or no toxicity. For some chemicals (e.g. aneugens) the toxicity curve may be very steep and require very small dose intervals to be used. If no toxicity is observed and the substance is freely soluble, the highest dose should correspond to at least 10 mM, 5 mg/ml or 5 µl/ml. For poorly soluble compounds, the highest dose should be above the limit of solubility in the final treatment medium at the end of the treatment. The middle and low doses should not contain precipitate. In most cases, this will mean doses should be separated by intervals of no more than $\sqrt{10}$.

- Controls

- **negative**

Since the solvent for each chemical should be specified, there is only the need to include solvent controls as negative controls. For additional experience, laboratories may include untreated controls as well. Negative control micronucleus (MN) frequencies for the various cell types should not exceed the following values (per 1000 cells):

	Without Cyto B	*With Cyto B*
CHO	35	35
CHL	35	35
V79	40	40
L5178Y	35	40
Human lymphocytes	N/A	35

N/A : not analyzed

- **positive**

Although many of the chemicals selected for the trial are known clastogens/aneugens, it is necessary to include positive controls for additional reference. This will allow to assess the quality of the preparations and to serve as a control to determine that the assay had functioned properly (e.g. S9 mix could activate an indirect genotoxin). Reference positive controls also allow determination of consistency of response from day to day.

- **performance**

The performance of positive control chemicals will be judged on the same basis as test chemicals (see "Criteria for positive call", below).

- Treatment schedules

- **cell lines**

<u>*Short term treatment*</u>: 3 – 6 hours in the presence and absence of S-9 followed by a period of treatment-free growth. Cells are sampled at a time equivalent to about 2 times the normal (e.g. untreated) cell cycle lengths after the beginning of treatment.

<u>*Continuous treatment*</u>: If negative or equivocal results are obtained, they should be confirmed using continuous treatment, or modified conditions as appropriate. In the study without metabolic activation, cells are exposed continuously for 2 – 2.5 cell cycles and then sampled.

- **lymphocytes**

Due to the fact that for the detection of MN *in vitro* in human lymphocytes Cytochalasin B is needed, the standard treatment schedule for CAT can not be applied.

The initial experiment would usually be in the absence of metabolic activation. Exposure to the test compound starts 24 hours after PHA stimulation, with treatment for 20 hours followed by the addition of Cytochalasin B at 44 hours and harvest at 72 hours after the beginning of the culture. If negative or equivocal results are obtained, a similar experiment is carried out but with start of treatment 48 hours after PHA stimulation.

If the protocols without metabolic activation give negative or equivocal result, a short exposure to the test compound (3–6 hours) in the presence of S9 is carried out, followed by washing to remove the compound. Cytochalasin B is added 44 hours after the start of the culture. In the case of negative or equivocal results, a similar experiment should be performed with S9 but with exposure for 3–6 hours at 48 hours after the PHA stimulation.

- No. of cultures
At least 2 replicate cultures per test concentration.

- No. of repeat experiments
At least 2 experiments for each treatment condition should be performed.
There is no requirement for verification of a clear-cut positive response. If negative the repeat experiment should be repeated following alternative treatment schedule(s).

- Analysis of slides
All slides, including those of positive and negative controls, should be independently coded before the microscopic analysis.
At least 2000 cells (mononucleate, or binucleate when cytochalasin B is used) should be scored for each test concentration. MN should demonstrate staining characteristics similar to that of the main nucleus, have a clearly defined membrane, be separate from the main nucleus and have a diameter no greater than 1/3 the diameter of the main nucleus.

- Criteria for positive call
There are several criteria for a positive response:
- dose-related increase in MN frequency (e.g. trend test)
- exceeds upper limit of historical negative control (see table above)
- statistically significant from concurrent control (chi-square or Fischer's exact test) OR fold increase (at least 2-fold) over control mean MN frequency

If a chemical satisfies all these criteria, it can be considered clearly positive. If a chemical satisfies none of these criteria, it can be considered clearly negative. If some, but not all of the criteria are met, the result will be inconclusive and further experimentation should be performed, probably with modified conditions (treatment/sampling time, dose intervals, level of toxicity, number of cells scored etc.).

STUDIES EVALUATED

Several studies (von der Hude et al., 2000; Miller et al., 1997; Miller et al., 1998; Matsumisha et al., 1999; Kirsch-Volder et al., 2003; Kirkland et al., 2005; Lorge et al., Aardema et al., Clare et al., Oliver et al., Wakata A et al., submitted for publication [Annexes 2-8]) were discussed and evaluated by the Expert Group during the first ECVAM meeting held in April 2004. The analysis was mainly based on the criteria for protocol requirements defined by the Group.
In the end, two data sets, the German Ring Trial (von der Hude et al., 2000; see Annex 3) and the French Ring Trial (Lorge et al.; submitted, see Annex 4) were considered to meet most, but not all, of the above set criteria for the ECVAM retrospective validation. All other studies cited above were considered in the assessment of the concordance between MNT *in vitro* and CAT *in vitro* (module 5) only and were used to support/strengthen or negate the conclusions drawn by the Validation Management Team (VMT).

a) German Ring Trial
von der Hude W. et al. (2000) In vitro micronucleus assay with Chinese hamster V79 cells – results of a collaborative study with in situ exposure to 26 chemical substances, Mutation Research 468 (2), 137-163 (see annex 3)

Organization of the ring trial

Ten laboratories participated in the collaborative study. The laboratory of the BgVV co-ordinated the work and delivered V79-cells and coded test substances. The S9-fraction from Aroclor 1254-induced rat liver was delivered by D. Utesch, Merck.

A detailed standard operation procedure (SOP) was developed. The laboratories were instructed about the solvent to be used for each substance and if the substance was to be tested without or with S9-mix only, or without and with S9-mix. The decision about the mandatory use of S9-mix was taken on the basis of available published results with the chromosomal aberration test *in vitro*.

In the first phase of the trial, all ten laboratories tested three coded substances (Griseofulvin, DMBA (+S9), and Pyrene). After the availability of the first results, the SOP was critically discussed and revised where necessary. The revised SOP was used for the second phase of the study.

During the second phase, each compound (23 in total) was tested by at least three laboratories and all results were sent to the co-ordinators. The analysis and discussion of the results were carried out once all results were available.

Selection of compounds:
The compounds tested, the solvents to be used and the delivering manufacturer are listed in Table 1. Chemicals were coded and sent to the participating laboratories.

Table 1: List of tested substances in the German Ring Trial

No	SUBSTANCE	CAS-No	SOLVENT	SOURCE*	S9-Mix
a) Aneugens					
1	Griseofulvin	126-07-8	DMSO	M	-
2	Diazepam	439-14-5	DMSO	S	-
3	Methyl-2-benzimidazole, Carbendazim	10605-21-7	DMSO	S	-
4	Thiabendazole	148-79-8	DMSO	S	-
5	Vinblastine	143-67-9	DMSO	S	-
6	Diethylstilbestrol, DES	56-53-1	DMSO	S	-
7	Hydroquinone	123-31-9	Ethanol	S	-
8	1,1,2-Trichloroethylene	79-01-6	DMSO	S	+/-
b) Indirect clastogens					
9	7,12-Dimethylbenzanthracene, DMBA	57-97-6	DMSO	S	+
10	2-Acetylaminofluorene	53-96-3	DMSO	S	+
11	Benzene	71-43-2	DMSO	M	+
c) Direct clastogens					
12	Actinomycin D	50-76-0	Ethanol	S	-
13	Cytosine arabinoside	147-94-4	dest. water	S	-
14	Bleomycin sulphate	11056-06-7	dest. water	S	-
15	Cadmium sulphate	10124-36-4	dest. water	S	-
16	Ethylmethanesulfonate, EMS	62-50-0	dest. water	S	-
17	5-Fluorouracil	51-21-8	dest. water	S	-
18	Methotrexate	59-05-2	DMSO	S	-
19	Mitomycin C, MMC	50-07-7	dest. water	S	-
20	Urethane	51-79-6	dest. water	S	
d) Non-mutagens					
21	Pyrene	129-00-0	DMSO	M	-
22	Benzylacetate	140-11-4	DMSO	S	-
23	Diethylhexylphthalate, DEHP	117-81-7	DMSO	S	+/-
24	Methylurea	598-50-5	dest. water	S	-
25	Toluene	108-88-3	DMSO	M	+/-
26	Retinol acetate	127-47-9	DMSO	S	-
Positive Controls used					
	Colchicine, COL °	477-30-5	dest. water	B	-
	Cyclophosphamide, CP °	50-18-0	dest. water	S	+

* B: Biochrom, Berlin, Germany; M: Merck, Darmstadt, Germany; S: Sigma-Aldrich, Deisenhofen, Germany
° Used as positive controls in each performed experiment

The overview of the study design and the protocols used is shown in table 3. The detailed test procedure is reported in the manuscript by von der Hude et al. (Annex 3).

b) SFTG Ring Trial

Lorge E. et al.; Aardema et al.; Clare et al.; Oliver et al.; Wakata A et al. SFTG international collaborative study on in vitro micronucleus test . Special Issue of Mutation Research. (submitted) (see Annex 4)

Organization of the ring trial

A total of 38 laboratories participated in the collaborative study. SFTG co-ordinated the work and delivered coded test substances. The objectives of the study were to evaluate different treatment protocols and the response of different cell systems. Detailed common protocols were developed, based on practices in use defined after a survey on the procedures used in the participating laboratories.

Four different cell types were used:

Number of labs	Cell lines used
10	Human Lymphocyte
8	CHO
14	CHL
6	L5178Y (mouse lymphoma)

This study aimed at evaluating different treatment-recovery schedules and conditions (see table 3), namely in the presence or absence of cytochalasin B. Therefore, no experiment was conducted with a metabolic activation system, in order to minimize the sources of variability. In addition, the use of a metabolic activation system was not expected to bring additional information on suitable treatment-recovery conditions.

Each compound was tested independently in two or three laboratories. At least two experiments were performed. The positive control was common to all the laboratories.

All results were collected on a standard template, sent to the co-ordinators and discussed when all results were available. The comparisons were based on the capacity of each treatment-recovery condition to detect the compound as positive or negative.

Selection of compounds:

The test substances, solvents to be used and the delivering manufacturer are listed in table 2. Chemicals were coded and sent to the test laboratory with the necessary instructions. The participants were instructed to handle each substance with precaution as if it is mutagenic/carcinogenic. The laboratories were instructed about the solvent to be used for each substance. No strict quantitative comparisons were made, as the compounds were tested blindly and therefore no determination of the absolute lowest effective concentration was performed.

The chemicals tested were chosen as representative of various modes of action and included non-genotoxic compounds. They were also chosen with regard to availability of results from other genotoxicity tests, especially *in vitro* chromosome aberration tests, *in vitro* micronucleus tests and *in vitro* mammalian cell gene mutation tests.

Table 2 List of tested substances in the SFTG study

SUBSTANCE	CAS-No	SOLVENT	SOURCE
Aneugens			
Griseofulvin	126-07-8	DMSO	Sigma
Thiabendazole	148-79-8	Water	Sigma
Diethylstilbestrol, DES	56-53-1	Ethanol	Sigma
Colchicine	64-86-8	Water	Sigma
Clastogens			
Cytosine arabinoside	147-94-4	Water	Sigma
Bleomycin sulphate	11056-06-7	Water	Sigma
5-Fluorouracil	51-21-8	Water	Sigma
Mitomycin C, MMC*	50-07-7	Water	Sigma
Urethane	51-79-6	Water	Sigma
Non-genotoxic compounds			
D-mannitol	69-65-8	Water	Sigma
Clofibrate	637-07-0	DMSO	Sigma

*: Used as the positive control in each experiment

The overview of the study design and the protocols used is shown in table 3. The detailed test procedure is reported in Annex 4.

The following table summarizes the main characteristics and differences of the two studies.

Table 3 Summary of the study designs and protocols used in the two studies.

Criteria	German Ring Trial	SFTG Trial
N. laboratories	10	38
SOP available	Yes	No
Cells	V 79	CHO (8 labs) CHL (14 labs) L5178Y (6 labs) human lymphocytes (10 labs)
Metabolic Activation Rat Liver S9 (Quality determined)	Yes, Merck	No
Test substances	9 direct clastogens, 3 indirect clastogens, 8 aneugens,6 non mutagens (Table 1)	4 clastogens, 4 aneugens, 2 non-genotoxic compounds (Table 2)
Source	BgVV coordination	SFTG coordination
Preparation Instructions provided by trail co-ordinator	Yes	Yes
Coded/Blinded Compound coded: (not necessary) Slides coded before analysis	 Yes Yes	 Yes Yes
Solvent/Vehicles Chemical solvent defined, where applic.	 Yes	 Yes
Cytochalasin B	No	All treatments performed with and without Cytochalasin B
Exposure conc/toxicity Measured relative to control * For details see annexes 3 and 4	PI index and/or MI (1000 cells)	With Cytochalasin B: % of multinucleated cells compared to solvent control. Without Cytochalasin B: cell count (1000 cells)
Adequate Controls Negative Positive	 Solvent Colcemid (without S9) Cyclophosphamide (with S9)	 Solvent Mitomycin C
Treatments Without Cytochalasin B With Cytochalasin B	 1) 3h treatment, 21h recovery (with S9) (S-S) 2) 24h treatment, no recovery (L-N) (without S9) No	 1) 3h treatment, 21h recovery (S-S) 2) 3h treatment, 45h recovery (S-L) 3) 24h treatment, no recovery (L-N) 4) 24h treatment, 24h recovery (L-L) 1) 3h treatment, 18-26h recovery (S-S) 2) 3h treatment, 45h recovery only HL (S-L)

		3) 24h treatment, 18-20h recovery; for HL 20h treatment, 28h recovery (L-L)
No. of Cultures- at least 2 replicates No. of Repeat Experiments- at least 2 experiments for each test condition	Yes (*in situ* method) Yes	Yes
Criteria for acceptability of the assay Statistically significant increase of MN in positive control as compared to solvent control At least one concentration between 50 and 60% At least 4 concentrations per genotoxicity assessment in at least one assay	 Yes Yes Yes	 Yes Yes Yes
Evaluation of micronuclei No of cells	 Min 1000 cells	 1000 cells/culture (2000 cells/concentration)
Criteria for positive call Dose-related increase in MN frequency Exceeds upper limit of historical controls Statistically significant from control (Chi-Sq./Fisher) Or 2 fold increase over control mean MN frequency Reproducibility of effects	 Yes No No No Yes	 Yes No Yes No Yes
Phase I	10 labs Griseofulvin - Aneugen DMBA + S9 - indirect clastogen Pyrene - non-mutagen	Only 1 phase
Phase II	10 labs 23 substances tested in at least 3 labs: full set of chemicals	
Applicability Domain	See Table 1	See Table2

EVALUATION OF THE STUDIES BY THE VALIDATION MANAGEMENT TEAM (VMT)

The careful evaluation of the two studies by the VMT led to the following considerations:

1) The scope and study designs of the two trials were different.
2) A huge number of variables were taken into consideration, especially in the SFTG trial (different cell lines, protocols, etc.).
3) The raw data of the two studies were evaluated by different expert groups.
4) The criteria considered for a positive call were not the same. In the German trial, biological relevance, a concentration-related increase of the MN frequencies and reproducibility of effects were the primary criteria for a positive call. In the SFTG study, the primary criteria were a concentration-related increase of MN frequency and a statistically significant increase in the incidence of micronucleated cells over the solvent control were considered.

Taking into account the above factors, it was clear that the data set was quite heterogeneous making it difficult to compare the data between studies. For this reason, in order to acquire more confidence in the data the VMT considered it necessary to re-analyse the raw data of both studies. The use of identical evaluation criteria led to a consistent call for both sets of raw data, allowing an improved final evaluation of the results.

The raw data from the German trial were provided to ECVAM by Silvio Albertini (Hoffmann-La Roche) and the raw data the SFTG trial were provided by Azeddine Elhaiouji (Novartis Pharma AG), who is the Editor of the special issue of Mutation Research on MNT SFTG trial. A series of manuscripts from the SFTG trial have been submitted to Mutation Research: one for each cell line and a general manuscript on the overall conclusions of the study.

The expert analysis of the raw data was conducted at ECVAM on the 14^{th}-15^{th} June 2005 during the Carcinogenicity Taskforce meeting. Four experts, which are also part of the ECVAM carcinogenicity taskforce, participated to the analysis of raw data: Hans-Juergen Ahr (Bayer HealthCare AG), Stefan Pfuhler (Wella, P&G), Jan van Benthem (RIVM, National Institute of Public Health and the Environment) and Philippe Vanparys (Johnson & Johnson). A consensus on the criteria for a positive call was reached among the experts prior to the evaluation of the raw data. The criteria were determined by taking into account: 1) the criteria initially defined by the expert group as if they had to be applied, in a best scenario case (see page 8); 2) the criteria defined in the draft OECD Test Guideline on the *in vitro* MNT (TG 487); and 3) the raw data available.

The tables summarizing all data re-evaluated by the VMT (tables 4, 5, 8, 10 and 18) were compiled by Marlies de Boeck and Natalie Mesens (Johnson & Johnson).

Criteria for the evaluation of raw data and the judgement of the relevance of effects

At the first meeting, the expert group agreed on a series of evaluation criteria as if they were for a prospective validation exercise. However, for this retrospective validation exercise, all criteria could not be applied in every case. Consequently, the criteria were overruled by an independent expert judgment, if considered necessary.

Statistical significance was not considered because it was not available for both studies. A judgement of the biological relevance of the effects observed was applied as the criterion to evaluate the data. The measure to assess the biological relevance of effects was the occurrence of a dose relationship and the magnitude of the effects. Historical control data were not available for the studies, which made it difficult to judge the relevance of relative increases compared to control. However, the observed range of the negative controls for each laboratory in this series of experiments was used as an aid to judge the relevance of effects. Increases of up to 6-fold, a value chosen by the experts, were considered irrelevant if they were due to very low control levels and if there was no dose response.

Definition of results being "equivocal"
If the use of the above described criteria did not allow to judge the individual experiment in question as positive, but the magnitude of the effect or the observed dose-relationship questioned the classification of the test item as negative, the study was rated equivocal.

Definition of results being "not appropriate" (na)
If in a study the required level of toxicity (50% or 60%) was not reached and no positive response was obtained, the study was rated as "NA". Rationale: It cannot be excluded that at a higher level of toxicity a positive result would have been obtained.

Additional information
In the SFTG study the judgment was based on binucleated cells, if results in both binucleated and mononucleated cells were available.
As in the German trial data on both proliferation index and mitotic index were not consistently available, both parameters were considered equally adequate for the determination of cytotoxicity. Rationale: based on the used *in situ* cultivation method the proliferation inhibition (toxicity) can be easily determined. Counting of 200 cells and determination of number of clones (1 cell, 2 cells, 3-4 cells, >5 cells) allows to calculate a proliferation index PI (for details see publication).

Table 4 provides an overview of the treatments and recovery times used in the two studies.

Table 5 reports the expert conclusions on the raw data from both studies that were selected for the retrospective analysis.

Table 6 gives an overview of the number of experiments, which were not appropriate according to the defined criteria.

Table 4: Overview of the treatments and recovery times used in the two studies

	SFTG Ring Trial							German Ring Trial	
	without CB				with CB			without CB	
Treatment →	S	S	L	L	S	S	L	L	S + S9
Recovery →	S	L	N	L	S	L	L	N	S
HL	nt	nt	nt	nt	3+26	3+45	20+28	nt	nt
L5178Y	3+21	3+45	24+0	24+24	3+20	nt	24+20	nt	nt
CHL	3+21	3+45	24+0	24+24	3+18	nt	24+18	nt	nt
CHO	3+21	3+45	24+0	24+24	3+20	nt	24+20	nt	nt
V79	nt	nt	nt	nt	nt	nt	nt	24+0	3+21

Abbreviations used:
S: Short
L: Long
N: No
CB: Cytochalasin B
nt: not tested
HL: Human Lymphocytes
L5178Y: Mouse Lymphoma Cells
CHL: Chinese Hamster Lung Cells
CHO: Chinese Hamster Ovarian Cells
V79: Chinese Hamster Lung Fibroblasts

Table 5: Overview of the within-laboratory variability

Table 6: not appropriate data

	SFTG Ring Trial							German Ring Trial	
	without CB				with CB			without CB	
Treatment →	S	S	L	L	S	S	L	L	S + S9
Recovery →	S	L	N	L	S	L	L	N	S
HL					4:16	5:17	2:17		
L5178Y	5:19	7:18	4:17	3:17	0:7		0:8		
CHL	5:32	10:30	5:33	10:32	2:16		1:16		
CHO	0:10		1:15	1:9	1:11		0:11		
V79								20:85	8:25

S: Short
L: Long
N: No
CB: Cytochalasin B
nt: not tested
HL: Human Lymphocytes
L5178Y: Mouse Lymphoma Cells
CHL: Chinese Hamster Lung Cells
CHO: Chinese Hamster Ovarian Cells
V79: Chinese Hamster Lung Fibroblasts

MODULE 2 - WITHIN-LABORATORY VARIABILITY

The within-laboratory variability assessment was based on the expert re-evaluation of raw data which took into account the 60% cytotoxicity criterion.

The same experiment was conducted twice in most of the laboratories involved in the SFTG study and in some laboratories in the German study (in certain instances it was conducted up to 4 times), allowing for the within-laboratory reproducibility assessment. Table 5 gives a schematic representation of all data collected and analysed.

Table 7 shows the within-laboratory reproducibility which was calculated for each treatment protocol and each cell line used in identical and independent experiments conducted more than once in the same laboratory. When the evaluation was carried out for each cell model and treatment protocol, the within-laboratory reproducibility ranged from 67% to 100%. The lowest value is related to the cell line CHL for the "Long Long" treatment. The within-laboratory reproducibility assessed per treatment, independent from cell model, varied from 84% to 100% (shown in red); while the reproducibility per cell line, independent from treatment, varied from 89% to 100% (shown in blue).

Table 5 Schematic representation of all data collected and analysed. For the analysis of the within-laboratory variability, the not appropriate data were excluded, but the equivocal data were included.

Table 7 Summary of the within-laboratory reproducibility results. The table presents the number and the percentage of laboratories which gave reproducible results for each treatment and each cell system. Only the laboratories that conducted identical experiments at least two times were considered. The within-laboratory reproducibility assessed per treatment, independent from cell model, is shown in red. The within-laboratory reproducibility per cell line, independent from treatment is shown in blue.

Table 7: Within-laboratory variability (Exclusion of non appropriate data)

	SFTG Ring Trial							German Ring Trial		
	without CB				with CB			without CB		
	S	S	L	L	S	S	L	L	S + S9	
	S	L	N	L	S	L	L	N	S	
HL					9:9 (100%)	7:7 (100%)	6:6 (100%)			22:22 (100%)
L5178Y	4:5 (80%)	5:6 (83%)	5:5 (100%)	6:6 (100%)	4:4 (100%)		5:6 (83%)			29:32 (91%)
CHL	13:13 (100%)	8:9 (89%)	11:12 (92%)	6:9 (67%)	9:9 (100%)		10:12(83%)			57:64 (89%)
CHO	6:6 (100%)	5:5(100%)	4:4 (100%)	5:5 (100%)	6:7 (86%)		5:7 (71%)			31:34 (91%)
V79								12:12 (100%)	nd	12:12 (100%)
	23:24 (96%)	18:20 (90%)	20:21 (95%)	17:20 (85%)	27:28 (96%)	7:7 (100%)	26:31 (84%)	12:12 (100%)		

S: Short
L: Long
N: No
CB: Cytochalasin B
nt: not tested
HL: Human Lymphocytes
L5178Y: Mouse Lymphoma Cells
CHL: Chinese Hamster Lung Cells
CHO: Chinese Hamster Ovarian Cells
V79: Chinese Hamster Lung Fibroblasts

MODULE 3 - TRANSFERABILITY

General Aspects

In general, the proposed test method can easily be performed in a laboratory that is experienced in routine cell culture techniques.

No extraordinary facilities are required. General cell culture laboratory equipment and instruments are sufficient to perform the proposed test method. All supplies and reagents are readily available on the market.

As stressed in the defined MNT testing requirements, when human lymphocytes are used they should derive from non-smoking, young healthy donors.

Training

The MNT *in vitro* requires personnel trained for general cell biology and cell culture activities (e.g. aseptic operations). Such expertise is available in most if not all QC-laboratories.

The operator should, in particular, be trained in the scoring of micronuclei. However, the training requirements for a person to be competent in scoring the slides are much less rigorous for MNT than for metaphase analysis. Moreover, as there is no requirement to count the chromosomes in a metaphase preparation, nor to evaluate subtle chromatid and chromosome damage, but only to determine whether or not a cell contains a micronucleus, the preparations can be scored much more quickly.

In addition, the successful transferability of the MNT *in vitro* is demonstrated by the satisfactory results for the between-laboratory variability from the two studies evaluated (see below, Module 4).

MODULE 4 - BETWEEN-LABORATORY VARIABILITY

As in the case of the within-laboratory variability, the between-laboratory variability was based on the expert conclusion of the raw data re-evaluation (Table 5). As shown in table 5, the between-laboratory variability has been assessed taking into account the 60% cytotoxicity criterion.

Since most of the laboratories repeated the identical experiment more than one time, the following criteria were considered to come to a final conclusion per each laboratory. These were applied when the results of an identical experiment conducted in the same laboratory were not concordant.
Positive + equivocal → positive
Negative + equivocal → negative
Positive + negative → equivocal

Table 8 gives a schematic overview of the between-laboratory variability for the different cell lines and the different treatments.

The data on the between-laboratory reproducibility per treatment protocol and per cell system are reported in table 9. In table 9, the not appropriate, inconclusive and equivocal data were excluded.

The between-laboratory reproducibility assessed per treatment, independent from cell line varied between 86% (for "Long Long" treatment) to 100%. The between-lab reproducibility assessed per cell model, independent from treatment, varies from 73% (for L5178Y) to 100%. Overall, taking into account all cell models and the different treatment, the between-laboratory reproducibility was 93% (93/100).

No major change in the between-laboratory variability was observed regarding between-laboratory reproducibility of the data in the case that the not appropriate data were excluded, while both the inconclusive and equivocal data were included in the analysis.

Table 8 Schematic representation of the between-laboratory variability.

Table 9 Summary of the between-laboratory reproducibility results. The table presents the number and the percentage of laboratories which gave reproducible results for each treatment and each cell system, taking into account the different chemicals analysed. The data reported refer to the experiments that have been conducted in at least two laboratories. Only the laboratories that conducted identical experiments at least two times were considered. The between-laboratory reproducibility assessed per treatment, independent from the cell model is shown in red. The within-lab reproducibility per cell line, independent from the treatment schedule, is shown in blue. Not appropriate, inconclusive and equivocal data have been excluded in this analysis.

Table 9: Between-laboratory variability (Exclusion not appropriate data, inconclusive and equivocal)

	SFTG Ring Trial							German Ring Trial		
	without CB				with CB			without CB		
Treatm.	S	S	L	L	S	S	L	L	S + S9	
Recov.	S	L	N	L	S	L	L	N	S	
HL					3:3 (100%)	5:5 (100%)	5:5 (100%)			13:13 (100%)
L5178Y	1:2 (50%)	3:3 (100%)	2:3 (67%)	2:3 (67%)	1:2 (50%)		2:2 (100%)			11:15 (73%)
CHL	5:5 (100%)	4:4 (100%)	5:5 (100%)	2:2 (100%)	4:4 (100%)		5:6 (83%)			25:26 (96%)
CHO	5:5 (100%)	5:5 (100%)	4:4 (100%)	3:3 (100%)	4:4 (100%)		5:5 (100%)			26:26 (100%)
V79								16: 18 (89%)	2:2 (100%)	18:20 (90%)
	11:12 (92%)	12:12(100%)	11:12 (92%)	7:8 (86%)	12:13(92%)	5:5 (100%)	17:18 (94%)	16: 18 (89%)	2:2 (100%)	

S: Short
L: Long
N: No
CB: Cytochalasin B
nt: not tested
HL: Human Lymphocytes
L5178Y: Mouse Lymphoma Cells
CHL: Chinese Hamster Lung Cells
CHO: Chinese Hamster Ovarian Cells
V79: Chinese Hamster Lung Fibroblasts

MODULE 5 – PREDICTIVE CAPACITY (CONCORDANCE)

The purpose of this retrospective validation is to determine whether the MNT *in vitro* can be used as alternative to the CAT *in vitro*. Therefore, module 5 will refer to **concordance** between the two tests, and not to predictive capacity.

The assessment of concordance was based on the following studies and reviews of published data selected by the expert group and the Validation Management Team:
1) German Trial (von der Hude et al., 2000; Annex 3);
2) Miller et al., 1997 (Annex 5);
3) Japanese Ring Trial (Matsushima et al., 1999; Annex 6);
4) Miller et al., 1998 (Annex 7);
5) Kirkland et al., 2005 (Annex 8).

The main aspects of the different data sets are presented and discussed below.

The French study was not designed to address concordance aspects of the MNT *in vitro*. Therefore, the Validation Management Team (VMT) decided not to consider this study for the assessment of concordance. The limited number of compounds tested with each protocol and treatment was not sufficient to draw justified conclusions on concordance. However, the amount of work in that study gave an added value for the within- and between-laboratory variability.

1) German Ring Trial

von der Hude W, Kalweit S, Engelhardt G, McKiernan S, Kasper P, Slacik-Erben R, et al. In-vitro micronucleus assay with Chinese Hamster V79 cells: results of a collaborative study with 26 chemicals. Mutation Research **468** 137 – 63 (2000). *(Annex 3)*

The German results could be considered with a higher degree of confidence since they were derived from the expert re-evaluation of the raw data.

Since the same experiment was repeated in several laboratories, the following criteria were used to come to a final conclusion per each substance. These criteria were applied when the results from an identical experiment that was conducted in different laboratories were not concordant.
Positive + equivocal → positive
Negative + equivocal → negative
Positive + negative → equivocal

Table 10 gives a schematic overview of the *in vitro* MNT results obtained using the above criteria.

Table 10 Overview of the conclusions from the VMT on the *in vitro* MNT in the German trial and the CAT reference data.

Reference data on CAT in vitro

Most of the references on the chromosome aberration test have been retrieved from the data set published by Kirkland et al. (2005). This review represents a huge database of over 700 chemicals that was compiled from different sources. To categorize the performance of the assays, Kirkland and his co-authors have re-evaluated the original data according to specific acceptability criteria described in the cited publication. For some of the compounds used in the German and SFTG studies, no data for the CAT were available in the Kirkland database. In such cases the references published by von der Hude et al. were considered.

A 2x2 contingency table was constructed (Table 11) for the results of the German trial from which the estimated concordance, specificity and sensitivity can easily be derived.

Table 11 Concordance between *in vitro* MNT and CAT of 24 compounds that gave clearly positive or negative results

		CAT results		Total
		+	-	
MNT results	+	12	4**	16
	-	2*	6	8
	Total	14	10	24

The not appropriate, equivocal (Methotrexate) and inconclusive (5-Fluorouracil) MNT data were not considered.

* Benzen, Urethane
** Methyl-2-Benzimidazole (Carbendazim), Diazepam; Thiabendazole, Retinal Acetate
(in red: established Aneugens)

Of the 4 chemicals which resulted positive in MNT and negative in CAT, 3 (Methyl-2-benzimidazole, Diazepam and Thiabendazole) are recognised aneugens, which we would expect to be positive in a MNT but may not be clastogenic in a CAT. Thus it could be considered that, in this study the *in vitro* MNT correctly predicted clastogenic or aneugenic status in 21/24 cases, i.e. a concordance of **87.5%**.

It should be noted that Urethane, which is classified positive in the CAT, resulted negative in the MNT in both German and SFTG trials. Although Urethane is classified as positive for CAT in the Kirkland database, it is also considered inconclusive in other published studies (Abe et al., 1977; Popescu et al., 1977).

In the German trial, no difference was observed between the analysis based on raw data evaluated considering the 50% cytotoxicity criterion and the ones evaluated with the 60% criterion.

Performance of MNT		
		Corrected for aneugens
Concordance MNT / CAT	**75%** (18/24)	**87.5%** (21/24)
Sensitivity	**85.7%** (12/14)	**88.2%** (15/17)
Specificity	**60%** (6/10)	**85.7%** (6/7)

2) Miller et al.

Miller B. et al. (1997) Comparative evaluation of the in vitro micronucleus test and the in vitro chromosomal aberration test: Industrial experience, Mutation Research 392,45-59 ; Albertini S. et al. (1997) Appendix: Detailed data on in vitro MNT and in vitro CAT: Industrial experience, Mutation Research 392,187-208 (Annex 5)

Four pharmaceutical companies evaluated the data from compounds tested in the *in vitro* CAT, as well as in the *in vitro* MNT. The compounds were tested either in Chinese hamster cell lines (CHO-K5, CHO-K1, V79) or in human peripheral blood lymphocytes. A total of 57 compounds were included in the analysis. However, the inconclusive compounds for MNT (compound 48), and for CAT (compounds 44 and 50) were not considered in the contingency table.

The strength of this data set is due to the fact that the compounds were tested in both assays with well established protocols (SOPs) and in parallel with the same cell line.

Results

Table 12 summarizes the concordance between MNT and CAT *in vitro*. A discussion on a compound by compound basis can be found in the original paper.

Table 12 Concordance between *in vitro* MNT and CAT with 57 compounds used in a comparative data evaluation of four pharmaceutical companies

		CAT results		
		+	-	Total
MNT results	+	19	8*	27
	-	1**	26	27
	Total	20	34	54

* Compounds 3, 5, **8**, **20**, 21, 24, **29**, **31** (8, 20 and 29 induced polyploidy and endoreduplication of chromosomes, 31 is a spindle poison)
** Compound 39

Of the 8 chemicals which resulted clearly positive in MNT and clearly negative in CAT, three (compounds 8, 20, and 29) are recognised to induce polyploidy and one (compound 31) was recognized as a spindle poison, according to the authors. Since it is expected that these chemicals may be positive in a MNT but may not be picked up in a CAT, it could be considered that, in this study the *in vitro* MNT correctly predicted clastogenic or aneugenic status in 49/54 cases, or in **90.7%** of occasions.

Performance of MNT		
		Corrected for aneugens
Concordance MNT / CAT	**83.3%** (45/54)	**90.7%** (49/54)
Sensitivity	**95%** (19/20)	**95.8%** (23/24)
Specificity	**76.5%** (26/34)	**86.7%** (26/30)

3) Japanese Ring Trial → CHL cells
Matsushima T., et al. (1999) Validation study of the in vitro micronucleus test in a Chinese hamster lung cell line (CHL/IU), Mutagenesis 14, 569-580 (Annex 6)

The Chinese hamster lung cell line CHL/IU was used to evaluate whether the *in vitro* MNT could be used as an alternative to the *in vitro* CAT. A total of 66 chemicals, including clastogens and polyploidy-inducers were evaluated. Treatments were carried out for 24, 48 or 72 hours in the absence of S9 mix, and/or for 6 hours with and without S9 mix followed by 18, 42 or 66 hours recovery. All chemicals were treated without using the Cytochalasin B cytokinesis block (CB) method and 1000 interphase cells were scored per dose level from at least 3 dose levels per treatment protocol. Additionally 5 chemicals were tested using the CB method and 1000 binucleate cells were scored per dose level. There was no enhancement in the ability to detect MN by using the CB method with these 5 chemicals.

Table 13 Concordance between *in vitro* MNT and CAT of 62 compounds evaluated

		CAT results		
		+	-	Totals
MNT results	+	43	7*	50
	-	4	8**	12
Totals		47	15	62

* Colchicine, Diethylstilbestrol, 4,4'-Methylene-bis(2-chloroaniline), *m*-Nitrotoluene, *o*-Nitrotoluene, Vinblastine Sulfate
** *p*-Nitrotoluene, 2-Methyl-4-nitroaniline

For 4 out of the 66 chemicals, no CA data were available. Therefore, the contingency table included a total of 62 chemicals.

In the CAT, compounds that did not induce structural chromosome aberration, but only induced numerical aberrations were considered negative. Among the positive MNT compounds that were negative for CAT *in vitro*, 6 compounds induced numerical aberrations (Colchicine, Diethylstilbestrol, 4,4'-Methylenebis (2-chloroaniline), *m*-Nitrotoluene, *o*-Nitrotoluene, Vinblastine Sulfate). Among the negative compounds in the MNT, four chemicals were positive for CAT (*p*-Chloroaniline, 2-Chloro-4-Nitroaniline, *o*-Nitroaniline, and Phenacetin) and two

compounds induced numerical aberrations (*p*-Nitrotoluene, 2-Methy-4-Nitroaniline). A possible explanation for the failure to induce MN by these 6 chemicals were given by the authors and mainly concerned differences in treatment conditions e.g. duration of treatment, top concentrations, spacing of doses. When all these factors were taken into account, the overall concordance between the MNT and CAT was calculated to be **88.7%** (55/62).

Performance of MNT		
		Corrected for inducers of numerical aberrations
Concordance MNT / CAT	**82.3%** (51/62)	**88.7%** (55/62)
Sensitivity	**91.5%** (43/47)	**89.1%** (49/55)
Specificity	**53.3%** (8/15)	**85.7%** (6/7)

4) GUM* Working Group → Literature Review

Miller B., et al. (1998) Evaluation of the in vitro micronucleus test as an alternative to the in vitro chromosomal aberration assay: position of the GUM working group on the in vitro micronucleus test, Mutation Research 410, 81-116 (Annex 7)

* GUM: Gesellschaft für Umweltmutationsforschung (German speaking Section of the European Environmental Mutagen Society EEMS)

A GUM working group performed an in-depth literature review to compare the *in vitro* MNT and CAT data to assess if the *in vitro* MNT can be used as an alternative/replacement of the *in vitro* CAT.

Selection of compounds/criteria for acceptance of publications

The initial selection of chemicals for evaluation by the GUM working group was based on a literature search (medline) for compounds that had been tested in both the MNT and the CAT. This first list consisted of 75 chemicals. For these compounds, a more detailed literature search in several databases (e.g. medline, toxall, toxline, embase) and a preliminary evaluation of the literature obtained was carried out.

Following this, rejection criteria were established, and papers were not selected for final evaluation if they fell into one or more of the following categories:

- written in a language other than English
- abstracts only
- review articles with no data
- tests system other than mammalian cells; cell lines established from rare diseases; repair deficient cell lines; primary cells other than human lymphocytes or Syrian hamster embryo (SHE) primary cells. Finally, the evaluation was limited to the following cells: 3T3, Swiss albino mouse fibroblasts; CHL, Chinese hamster lung fibroblasts; CHO, Chinese hamster ovary cells; DON, Chinese hamster lung cells; HULY, human (peripheral blood)

lymphocytes; L5178Y, mouse lymphoma cells; SHE; V79, Chinese hamster lung fibroblasts; HepG2, human hepatocellular carcinoma cells

- method and results not explained in detail
- compound concentration not transferable to µg/ml
- no negative control given (although a control for only one sampling time or historical control data were accepted); a positive control was not required
- data given for only one concentration of the test compound
- number of cells analyzed lower than 100 (CAT) or 1000 (MNT) or not given
- no information about the kind of lesion in the CAT

If less than two acceptable MNT publications were available, the compound was eliminated from the list. CAT publications were not required at this point in order to avoid exclusion of aneugens from the database. No additional systematic literature search was carried out after the end of 1995.

Each individual publication was then evaluated according to the following criteria:

- type of assay (MNT or CAT) and cell type
- use of Cytochalasin B (in the MNT) and of S9 mix
- concentration range (µg/ml) from the lowest to the highest concentration applied
- treatment time and sampling time; both given as hours after start of treatment
- cytotoxicity endpoint (if sufficient information was provided).
- highest MN or CA frequency as percent of cells with micronuclei or aberrations (excluding gaps) in the most effective treatment protocol presented
- author's evaluation of the result as positive/negative/inconclusive (or the implication of a positive result by the author)
- evaluation by the working group as positive/negative/inconclusive according to the overall impression of the experimental result. A doubling over control was not necessarily considered adequate by itself. In the case of deviations from the author's evaluation, data were discussed by the working group.
- the lowest concentration, if considered by itself, that yielded a positive result, was given as lowest effective concentration tested (LOED; in µg/ml)
- if there were at least two consecutive concentrations having increased aberration frequencies (compared with the concurrent negative control) and the effect of the higher concentration was more pronounced than that of the lower, the effect was labelled as "DER yes" (dose-effect relationship)
- if at least two data sets (possibly with modified methodology) from the same cell line were shown, the result was considered to be confirmed in one publication
- acceptance of a publication in spite of variations from the above requirements, and further information regarded as important by the working group, yielded a remark:
  1) no. of cell evaluated not given; 2) frequency in no. of MN/CA per 100 cells (not % cells with MN/CA 3) high toxicity genotoxin; 4) frequency of chromosomal aberrations including gaps; 5) mitotic shake-off-method; 6) control level subtracted; 7) high concentration of solvent (e.g. 3.3% DMSO); 8) no concurrent control value given.

The final database obtained included 96 publications and covered 34 compounds. Only for 30 compounds data were available for both tests.

Table 14 List of all compounds reported to be tested in the *in vitro* MNT.

Compound (abbreviation)	***CAS number***
2-Acetylaminofluorene (2AAF)	53-96-3
Actinomycin D (AMD)	50-76-0
Adriamycin (ADR)	25316-40-9
Aflatoxin B1 (AFB1)	1162-65-8
2-Aminoanthracene (2AA)	613-13-8
m-Amsacrine (MAC)	54301-15-4
Benzo(a)pyrene (BaP)	50-32-8
Bleomycin sulfate (BLM)	11056-06-7
Cadmium chloride (CD)	10108-64-2
Chloralhydrate (CH)	302-17-0
Colchicine (COL)	64-86-8
Cyclophosphamide (CP)	50-18-0
Diazepam (DZ)	439-14-5
Diethylstilbestro (DES)	56-53-1
7,12-Dimethylbenzanthracene (DMBA)	57-97-6
Econazole (EZ)	27220-47-9
Ethylmethanesulfonate (EMS)	62-50-0
5-Fluorouracil (5FU)	51-21-8
Griseofulvin (GF)	126-07-8
Hydroquinone (HQ)	123-31-9
Methyl-2-Benzimidazole Carbamate (MBC)	10605-21-7
3-Methylcholanthrene (MCA)	56-49-5
Methylmethanesulfonate (MMS)	66-27-3
N-Methyl-N'-nitro-N-nitroso-guanidine (MNNG)	70-25-7
1-Methyl-1-Nitrosourea (MNU)	684-93-5
Mitomycin C (MMC)	50-07-7
Neocarcinostatin (NCS)	9014-02-2
2-Nitrofluorene (NF)	607-57-8
Phenol (PHE)	108-95-2
Pyrene (PYR)	129-00-0
Pyrimethamine (PY)	58-14-0
Thiabendazole (TB)	148-79-8
Thimerosal (TM)	54-64-8
Vincristine sulfate (VCR) (other salt were used in some publications)	5722-7

Colcemid and Vinblastine were not included, as no additional information to that obtained with the structurally related compounds, Colchicine and Vincristine, would have been gained.

Results

Table 15 summarizes the concordance between MNT and CAT *in vitro* based on the evaluation of the GUM working group. A discussion on a compound by compound basis is reported in the original paper.

Table 15 Concordance between *in vitro* MNT and CAT of 30 compounds for which data used in a literature review by a GUM Working Group

		CAT results		
		+	-	Total
MNT results	+	23	3*	26
	-	-	1	1
	Total	23	4	27

The inconclusive MNT data (Pyrimethamine, Thiabendazole) and inconclusive CA data (Griseofulvin) were not considered.

* Diazepam, Diethylstilbestrol, Methyl-2-Benzimidazol-Carbamate (known *in vitro* aneugens)

The three discordant compounds showing increases in the number of MN but no CA induction are all known or suspected aneugens. The detection of these compounds underlines the additional strength of the *in vitro* MNT. Thus it could be considered that, in this study the *in vitro* MNT correctly predicted clastogenic or aneugenic status in all cases, or in **100%** of occasions.

Performance of MNT		
		Corrected for aneugens
Concordance MNT / CAT	**88.9%** (24/27)	**100%** (27/27)
Sensitivity	**100%** (23/23)	**100%** (26/26)
Specificity	**25%** (1/4)	**100%** (1/1)

No conclusions can be drawn on specificity due to the low number of negative compounds considered in the literature review and their aneugenic properties.

5) Kirkland et al. review (2005)

Kirkland D, et al. 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. Mutation Research, 584 1-256 (Annex 8)

Kirkland *et al* (2005) reviewed the published genotoxicity results with more than 900 chemicals of defined carcinogenic or non-carcinogenic status in rodents. *In vitro* MNT and CAT results were available for many chemicals. For those that gave clearly positive or negative results, the following concordance was observed (table 16):

Table 16 Concordance between *in vitro* MNT and CAT of 88 compounds

		CAT results		
		+	-	Totals
MNT results	+	57	12*	69
	-	11	8	19
	Totals	68	20	88

* Diazepam, 17-β-estradiol, Oxazepam, Nitrilotriacetic acid and Rotenone (recognised aneugens)

Thus MNT and CAT results agreed with one another in 65/88 cases, i.e. for 73.9% of chemicals. Of the 12 chemicals that were negative in CAT but positive in MNT, 5 (Diazepam, 17-β-estradiol, Oxazepam, Nitrilotriacetic acid and Rotenone) are recognised aneugens, which we would expect to be positive in an MNT but may not be clastogenic in a CAT. Thus it could be considered that, in this database, the *in vitro* MNT correctly predicted clastogenic or aneugenic status in 70/88 cases, or **79.5%** of occasions. In addition, Trichloroethylene and Carbon Tetrachloride, which are negative in the CAT test but positive in the MNT, may also induce aneuploidy.

Because most of the published MNT and CAT results were from different laboratories at different times and the level of toxicity achieved in the CAT was not recorded and may have been high enough to result in artefactual positive results, this concordance is considered very satisfactory.

Performance of MNT		
		Corrected for aneugens
Concordance MNT / CAT	**73.9%** (65/88)	**79.5%** (70/88)
Sensitivity	**83.8%** (57/68)	**84.9%** (62/73)
Specificity	**40%** (8/20)	**53.3%** (8/15)

Overall Concordance

Different studies have been presented in this module. These studies differ one from the other due to several characteristics like the availability of raw data, whether or not the MNT and CAT were conducted in parallel within the same study, the quality of CAT reference data considered, the use of proprietary compounds and the number of compounds tested. The results are summarised in table 17. These features are critical since they confer a different weight to the studies. Moreover, it should be noted that some compounds have been tested in more than one study and that the review studies might include data that have already been reported in other published studies.

The concordance between *in vitro* MNT and *in vitro* CAT ranges from 73.9% to 88.9% in the different studies. If suspect or known aneugens were considered, the *in vitro* MNT correctly predicted clastogenic or aneugenic status in 79.5% to 100% of cases.

Table 17 summarizes the main features and the performance of the studies considered in module 5.

Table 17 Overview the main features and concordance data of the studies considered.

	MNT raw data	MNT & CAT	Reference CAT	Proprietary data	No compounds	Concordance %	Conc. corrected %	Sens. %	Sens. corrected %	Spec. %	Spec. corrected %
1) German ring trial	X		K DB		24	75	87.5	85.7	88.2	60	85.7
2) Miller et al., 1997		X		X	54	83.3	90.7	95	95.8	76.5	86.7
3) Japanese ring trial			literature		62	82.3	88.7	91.5	89.1	53.3	85.7
4) GUM Working group			literature		27	88.9	100	100	100	25	100
5) Kirkland et al., 2005			K DB		88	73.9	79.5	83.8	84.9	40	53.3

K DB: Kirkland et al. database, 2005

MODULE 6 – APPLICABILITY DOMAIN

Toxicological endpoints:
Structural and numerical chromosome aberration leading to the formation of micronuclei are the endpoints of genotoxicity.

Chemical classes:
The *in vitro* MNT can be used for clastogens, agents giving rise to structural chromosomal aberrations in cells, and aneugens, agents (e.g. spindle poisons) which cause changes in the number of chromosomes per cell (numerical chromosomal aberrations). As shown in the studies considered for the evaluation of the MNT (modules 2-5), the chemicals used in these studies covered a broad range of chemical classes.

Regulatory uses*:*
Based on the data analysed the MNT *in vitro* can be applied in human toxicology for chemicals as alternative to the CAT *in vitro*. The test has also the potential to be used for agrochemicals, pharmaceuticals and in the field of ecotoxicology, where it may be useful to study genotoxicity in fish (Al- Sabti and Metcalfe, 1995).

ADDITIONAL INFORMATION

Considerations related to the use of 50% and 60% cytotoxicity (analysis by the VMT)
The re-evaluation of the SFTG and German trial raw data by the VMT was carried out both for 50% and 60% cytotoxicity (Table 18). In general, not much difference was observed between the data evaluated considering acceptable a maximum dose of 50% and the ones based only on experiments where 60% cytotoxicity was reached. The conclusions drawn from the results generated with the 8 treatment protocols did not change when a different cytotoxicity level was applied. No changes from positive to negative results were observed, with the exception of Cytosine Arabinoside tested in human lymphocytes. However, Cytosine Arabinoside resulted clearly positive in the other cell lines tested. Five results considered positive at 60% cytotoxicity became inconclusive if 50% cytotoxicity was considered as sufficient (meaning that the value observed at 60% cytotoxicity was ignored because the test compound would not have been tested at such a high level of cytotoxicity). Two non appropriate and one equivocal results at 60% cytotoxicity became negative at 50%.

Table 18 Overview of the conclusions from the SFTG and German trials, considering 50% versus 60% cytotoxicity, using different protocols, different cell systems

CONCLUSION BY THE VALIDATION MANAGEMENT TEAM

The primary focus of the ECVAM 'retrospective' validation using the modular validation approach was an evaluation of the potential of the *in vitro* micronucleus test as alternative to the standard *in vitro* chromosome aberration assay.

In the past decade several studies comparing the *in vitro* chromosome aberration assay and the micronucleus *in vitro* assay were performed. A high correlation was observed (>85%) (von der Hude et al., 2000; Miller et al., 1997; Miller et al., 1998).

The working group evaluated in a first step available published data and came to the conclusion, that two publications met the criteria for a retrospective validation:

1. von der Hude W., et al. (2000) In vitro micronucleus assay with Chinese hamster V79 cells – results of a collaborative study with in situ exposure to 26 chemical substances, Mutation Research 468 (2), 137-163
2. Lorge E. et al. (2004) SFTG international collaborative study on *in vitro* micronucleus test I. General conditions and overall conclusions of the study. Submitted Mutation Research.

Additional published data were considered to address the concordance between the MNT *in vitro* and the CAT *in vitro* (module 5) and to confirm/strengthen the conclusions reached based on the above two data sets.

The earlier mentioned high correlation between the MNT and CAT was confirmed by the ECVAM Expert group. The concordance ranges were between 73.9% and 88.9% and between 79.5% and 100% (corrected for known aneugens) taking the CAT *in vitro* as 'Gold Standard'.
The observed values are in-line with other known and well accepted concordances. For instance, the concordance between carcinogenic response in rats and mice for chemicals tested in both species is about 75% (Gold and Zeiger, 1997).

Conclusions of the Validation Management Team for the different Modules

			Short Summary & Conclusion
Module 1	Test Definition	√	Clear definitions of the scientific basis, description of the endpoint(s) and the mechanistic basis; protocol requirements available.
Module 2	Within-Laboratory Variability	√	The within laboratory-variability was in an acceptable range (84% to 100% assessed per treatment independent from cell model; 89% to 100% reproducibility per cell line, independent from treatment)
Module 3	Transferability	√	Test method can be easily transferred and no extraordinary facilities are required. Overall, the successful transferability of the MNT *in vitro* is demonstrated by the satisfactory results for the between-laboratory variability from the two studies evaluated (Module 4).
Module 4	Between-Laboratory Variability	√	The between-laboratory reproducibility assessed per treatment, independent from cell line varied between 86% and 100%. The between-laboratory reproducibility assessed per cell model, independent from treatment varies from 73% to 100%.
Module 5	Predictive Capacity (Concordance)	√	The concordance between MNT *in vitro* and CA *in vitro* ranges from 73.9% to 88.9% in the different studies. Corrected for compounds known to induce aneuploidy (detected by the *in vitro* MNT) the range is 79.5 to 100%.
Module 6	Applicability Domain	√	Genotoxicity (structural and numerical chromosome aberration); all chemical classes; potential to be used in screening strategy for genotoxicity evaluation (regulatory use).
Module 7	Minimum Performance Standards		Not considered in this evaluation of retrospective data.

Potential as alternative to in vivo MNT

The potential as alternative to the *in vivo* micronucleus assay was not assessed in-depth by the group. However, there is supportive data/evidence, that the MNT *in vitro* has a high predictivity for the *in vivo* MNT.

Based on the data presented in this report, the Validation Management Team concluded that the MNT *in vitro* does fulfil the criteria for a successful validation and can be used as an alternative/replacement for the CAT *in vitro*.

REFERENCES

Aardema MJ. et al. SFTG International collaborative study on the in-vitro micronucleus test. III Using CHO cells. Mutation Research., submitted.

Abe S and Sasaki M. Chromosome aberrations and sister chromatid exchanges in Chinese hamster cells exposed to various chemicals. Journal of National Cancer Institute, **58**(6)1635-41 (1977).

Albertini S, Miller B, Chetelat AA, Locher F. Detailed data on the in-vitro micronucleus test and in-vitro chromosome aberrations; industrial experience. Mutation Research, **392** 45 – 59 (1997).

Al- Sabti K and Metcalfe CD. Fish micronuclei for assessing genotoxicity in water. Mutat Research **343** 121-35 (1995).

Clare G et al. SFTG International collaborative study on the in-vitro micronucleus test. II Using human lymphocytes. Mutation Research., submitted (2004).

de Stoppelaar JM et al. Analysis of DES-induced micronuclei in binucleated rat fibroblasts: comparison between FISH with a rat satellite I probe and immunocytochemical staining with CREST serum**.** Mutation Research, **392**(1-2) 139-49 (1997).

de Stoppelaar JM et al. Increased frequencies of diploid sperm detected by multicolour FISH after treatment of rats with carbendazim without micronucleus induction in peripheral blood erythrocytes. Mutagenesis, **14**(6) 621-32 (1999).

Eastmond D.A and Pinkel D. Detection of aneuploidy and aneuploidy inducing agents in human lymphocytes using fluorescence in-situ hybridisation with chromosome specific DNA probes. Mutation Research, **234** 9 – 20 (1990).

Eastmond D.A and Tucker J.D. Identification of aneuploidy inducing agents using cytokinesis-blocked human lymphocytes and an antikinetochore antibody. Environmental Molecular Mutagenesis, **13** 34 – 43 (1989).

Eastmond D.A et al. Detection of hyperploidy and chromosome breakage in interphase in human lymphocytes following exposure to the benzene metabolite hydroquinoine using multicolour fluorescence in-situ hybridisation with DNA probes. Mutation Research, **322** 9 – 20 (1994).

Farooqi Z et al. The use of fluorescence in-situ hybridisation for the detection of aneugens in cytokinesis- blocked mouse splenocytes. Mutagenesis, **8** 329 – 34 (1993).

Fenech M and Morley AA. Cytokinesis-block micronucleus method in human lymphocytes: effect of in-vivo ageing and low dose X-irradiation. Mutation Research, **161**, 193 – 8 (1986).

Galloway SM, Armstrong MJ, Reuben C, Colman S, Brown B, Cannon C, Bloom AD, Nakamura F, Ahmed M, Duk S, et al.Chromosome aberrations and sister chromatid exchanges in Chinese hamster ovary cells: evaluations of 108 chemicals. Environmental and Molecular Mutagenesis, **10** Suppl 10:1-175, (1987)

Garriott ML, Phelps JB, Hoffman WP. A protocol for the in-vitro micronucleus test I Contribution to the development of a protocol suitable for regulatory submissions from an examination of 16 chemicals with different mechanisms of action and level of activity. Mutation Research, **517** 123 – 34 (2002).

Gold LS and Zeiger E. Handbook of Carcinogenic Potency and Genotoxicity Databases. eds. Boca Raton FL: CRC Press (1997)

Gulati DK, Witt K, Anderson B, Zeiger E, Shelby MD.Chromosome aberration and sister chromatid exchange tests in Chinese hamster ovary cells in vitro. III. Results with 27 chemicals. Environmental and Molecular Mutagenesis, **13**(2) 133-93, (1989).

Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Hoffmann S, Roi AJ, Prieto P, Sabbioni E, Scott L, Worth A, Zuang V. A modular approach to the ECVAM principles on test validity. ATLA, **32**(5) 467-72 (2004).

Honma M, Hayashi M, Shimada H, Tanaka N, Wakuri S, Awogi T, Yamamoto KI, Kodani N, Nishi Y, Nakadate M, Sofuni T. Evaluation of the mouse lymphoma tk assay (microwell method) as an alternative to the in vitro chromosomal aberration test. Mutagenesis, **14**(1) 5-22, (1999).

Keshava C, Keshava N, Whong WZ, Nath J, Ong T. Inhibition of metho trexate-induced chromosomal damage by folinic acid in V79 cells. Mutation Research, 397 221-228 (1998)

Kirkland D, Aardema M, Henderson L and Müller L. 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. Mutation Research, **584** 1-256 (2005).

Kirsch-Volders M, Sofuni T, Aardema M, Albertini S, Eastmond D, Fenech M et al. Report from the in-vitro micronucleus assay working group. Mutation Research, **540** 153 – 63 (2003).

Kirsch-Volders, M. Towards a validation of the micronuclei test. Mutation Research, **392** 1-4 and contents of whole of this Special Issue (1997).

Ishidate M Jr, Harnois MC, Sofuni T. A comparative analysis of data on the clastogenicity of 951 chemical substances tested in mammalian cell cultures. Mutation Research, **195**(2) 151-213 (1988).

Lorge M.J et al. SFTG International Collaborative Study on in-vitro micronucleus test I. General conditions of the study. Mutation Research., submitted.

Matsushima T, Hayashi M, Matsuoka A, Ishidate M Jr, Miura KF, Shimizu H**,** et al. Validation study of the in-vitro micronucleus test in a Chinese hamster lung cell line (CHL/IU) Mutagenesis, **14** 569 – 80 (1999).

Matsuoka A, Yamakage K, Kusakabe H, Wakuri S, Asakura M, Noguchi T, Sugiyama T, Shimada H, Nakayama S, Kasahara Y, Takahashi Y, Miura KF, Hatanaka M, Ishidate M Jr, Morita T, Watanabe K, Hara M, Odawara K, Tanaka N, Hayashi M, Sofuni T**.** Re-evaluation of chromosomal aberration induction on nine mouse lymphoma assay "unique positive' NTP carcinogens. Mutation Research, **369**(3-4) 243-52 (1996).

Migliore L, Bocciardi R, Macri C, Lo Jacono F. Cytogenic damage induced in human lymphocytes by 4 vanadium compounds and micronucleus assay by fluorescence in-situ hybridisation with a centromeric probe. Mutation Research, **319** 205 – 13 (1993).

Miller B, Albertini S, Locher F, Thybaud V, Lorge E**.** Comparative evaluation of the in-vitro micronucleus test and the in-vitro chromosome aberrations test: industrial experience. Mutation Research, **392** 45 – 59 (1997).

Miller B, Potter-Locher F, Seelbach A, Stopper H, Utesch D, Madle S**.** Evaluation of the in-vitro micronucleus test as an alternative to the in-vitro chromosome aberrations assay: position of the GUM working group on the in-vitro micronucleus test. Mutation Research, **410** 81 – 116 (1998).

Miller BM, Zitzelsberger HF, Weier HU, Adler ID. Classification of micronuclei in murine erythrocytes: immunofluorescent staining using CREST antibodies compared to in-situ hybridisation with biotinylated satellite DNA. Mutagenesis**, 6** 297 – 302 (1991).

Norppa H, Renzi L, Lindholm C. Detection of whole chromosomes in micronuclei of cytokinesis-blocked human lymphocytes by antikinetochore staining and in-situ hybridisation. Mutagenesis, **8** 519 – 25 (1993).

OECD Test Guideline 473 (in vitro mammalian chromosome aberration test), revised February 1998.

Oliver et al. SFTG International collaborative study on the in-vitro micronucleus test. V. Using L5178Y cells. Mutation Research., submitted.

Parry J .M. and Sorrs A. The detection and assessment of the aneugenic potential of environmental chemicals: the European Community aneuploidy project. Mutation Research, **287** 3-15 (1993).

Popescu NC, Turnbull D, Di Paolo JA. Sister chromatid exchange and chromosome aberration analysis with the use of several carcinogens and non-carcinogens. Journal of National Cancer Institute, **59**(1) 289-93 (1977).

Sasaki MS. Chromosome aberration formation and sister1 chromatid exchange in relation to DNA repair in human cells. Basic Life Science, **15** 285-313, (1980).

Segawa M, Nadamitsu S, Kondo K, Yoshizaki I. Chromosomal aberrations of Don lung cells of Chinese hamster after exposure to vinblastine in vitro. Mutation Research, **66**(1) 99-102, (1979).

Technical Guidance Document on Risk Assessment, (in support of Commission Directive 93/67/EEC on risk assessment for new notified substances, Commission Regulation (EC) No 1488/94 on risk assessment for existing substances and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market), European Commission Joint Research Centre, EUR 20418 EN/1, part I (3.10), pp 144-154, (2003).

von der Hude W, Kalweit S, Engelhardt G, McKiernan S, Kasper P, Slacik-Erben R, et al. In-vitro micronucleus assay with Chinese Hamster V79 cells: results of a collaborative study with 26 chemicals. Mutation Research, **468** 137 – 63 (2000).

Wakata A et al. SFTG International collaborative study on the in vitro micronucleus test. IV. Using CHO/IU cells. Mutation Research, submitted.

ANNEXES

Annex 1. OECD Draft Test Guideline on MNT *in vitro*

Annex 2. IWGTP Report (Kirch-Volders et al., 2003)

Annex 3. Manuscript von der Hude et al. (2000)

Annex 4. Submitted manuscripts SFTG Ring Trial

Annex 5. Manuscript Miller et al. (1997)

Annex 6. Manuscript Japanese Ring Trial (Matsushima et al., 1999)

Annex 7. Manuscript GUM Working Group (Miller et al., 1998)

Annex 8. Manuscript Kirkland et al. (2005)