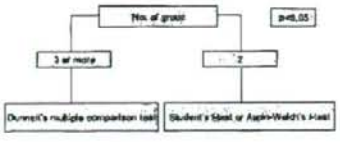
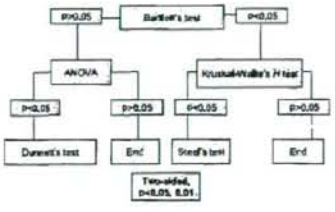
Statistical tools used in short-term toxicity studies.

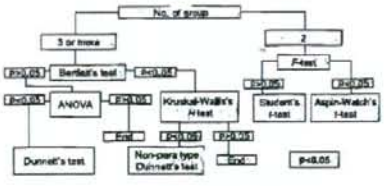
Tool No. 1, use rate;5/122



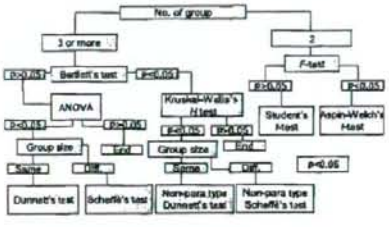
Tool No. 2, use rate;7/122



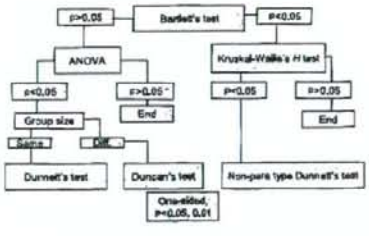
Tool No. 3, use rate;9/122



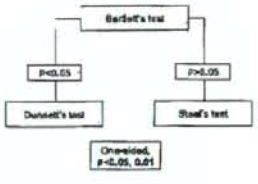
Tool No. 4, use rate;10/122



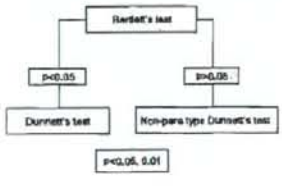
Tool No. 5, use rate;9/122



Tool No. 6, use rate;20/122



Tool No. 7, use rate;10/122



Tool No. 8, use rate;23/122

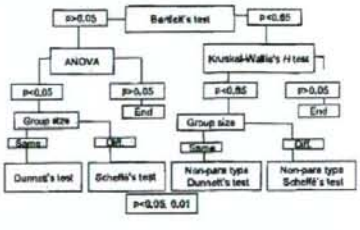


Fig. 1. Classification of number of studies based on the statistical tools used for the analysis of quantitative data.

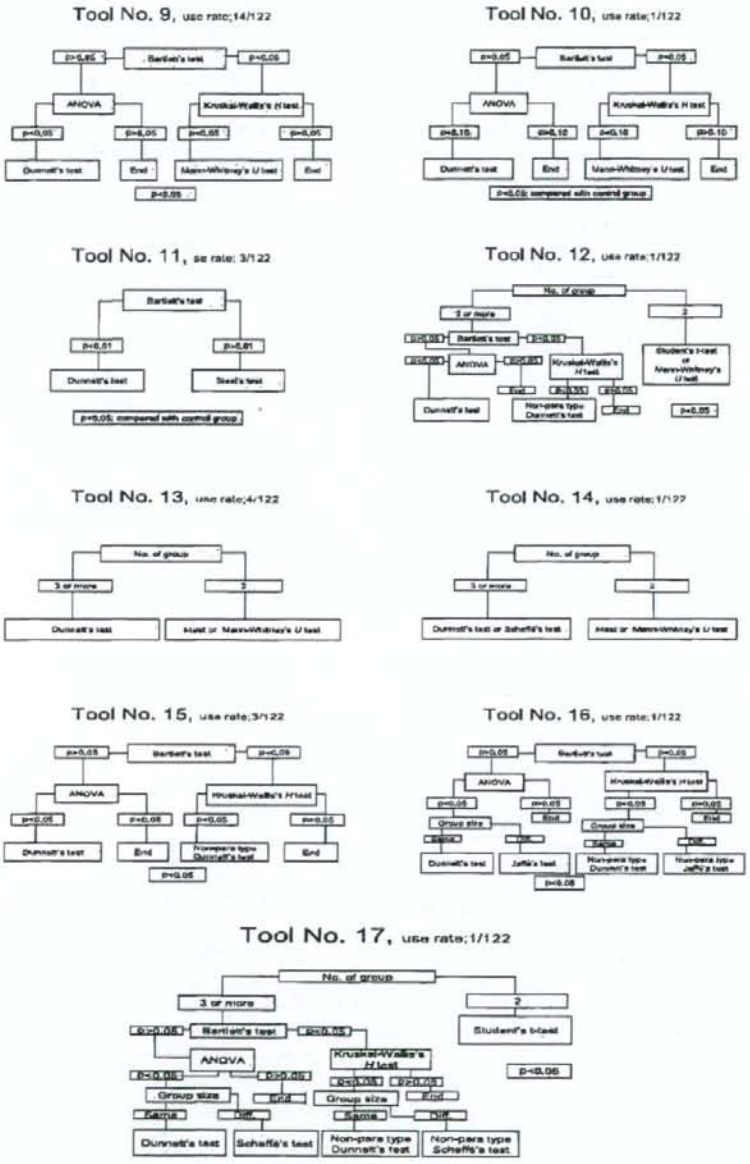


Fig. 1. Continued.

be noted that use of ANOVA causes the error of the second kind. Because of this, some of the recent studies skipped ANOVA in the decision tree and straight away used the statistical tools for *post hoc* comparison (Sumida *et al.*, 2006; Nagano *et al.*, 2006).

For the analysis of qualitative data, chi-square and Fisher's tests do not seem to be appropriate, though Fisher's test is slightly more sensitive than the chi-square test. These two tests do not detect a significant difference between a finding in the dosage group and control group, when all the animals (5/5) show the finding in the dosage group and 2 animals in the control group (2/5). On the other hand, Mann-Whitney's *U* test, which converts the scores into numerical values, detects a significant difference, when the finding in the dosage group is 5/5 and control group is 2/5. Therefore, Mann-Whitney's *U* test has better sensitivity to analyze qualitative data than the chi-square and Fisher's tests. Trend test like Jonckheere test can be used to determine no observed adverse effect level/ no observed effect level (NOAEL/NOAL) in the twenty-eight-day repeated dose oral toxicity tests. The statistical tools used, especially in the case of non-parametric tests, to determine the NOAEL/NOAL may be clearly elaborated in the study report.

We propose Dunnett's test for the analysis of quantitative data obtained from twenty-eight-day repeated dose oral toxicity tests in rodents and for qualitative data, Mann-Whitney's *U* test. For both tests, one-sided test with p=0.05 may be applied.

Table 2. A classification of number of studies based on the statistical tools used for the analysis of qualitative data.

Tool. No.	Description of statistical tools		Number of studies
1	Scored data	Frequency data	6
	Mann-Whitney's *U* test (two-sided, p<0.05)	Fisher's test (one-sided, p<0.05)	
2	Urinalysis	Pathological findings	7
	Cumulated Chi-square test (two-sided, p<0.05, p<0.01)	Mann-Whitney's *U* test (two-sided, p<0.05, p<0.01)	
3	Urinalysis	Pathological findings	13
	Cumulated Chi-square test (p<0.05)	Mann-Whitney's *U* test (two-sided, p<0.05)Fisher's test (one-sided test, p<0.05)	
4	Pathological findings		26
	Fisher's test (one-sided test, p<0.05)		
5	Pathological findings		19
	Chi-square test (p<0.05)		
6	FOB, urinalysis and differential leucocytes		15
	Kruskal-Wallis's *H* test, Mann-Whitney's *U* test (p<0.05)		
7	Urinalysis and pathological findings		9
	Mann-Whitney's *U* test (two-sided, p<0.05, p<0.01)		
8	Pathological findings		1
	Fisher's test		
9	FOB, sense function test and macroscopic and microscopic findings of pathology		1
	Wilcoxon rank-sum test, Fisher's test and Mann-Whitney's *U* test (p<0.05, p<0.01)		
10	Pathological findings		4
	Nonparametric type Dunnett's test or non-parametric type Scheffe's test, and Cochran-Armitage's trend test		
11	FOB, sense function test and macroscopic and microscopic findings of pathology		21
	No statistical tool mentioned		
Total			122

Table 3. Use of one-sided or two-sided test for short-term repeated dose administration toxicity studies with rats.

Data	One-sided	Two-sided	No mentioned	Total
Quantitative	22	13	87	122
Qualitative	34	22	70	126

ACKNOWLEDGMENT

Research described in this paper was supported by Grant (Project name: Development of Hazard Assessment Techniques Using Structure-activity Relationship Methods) from New Energy and Industrial Technology Development Organization (NEDO).

REFERENCES

Dunnett, C.W. (1964): New tables for multiple comparisons with a control. Biometrics, September, 482-491.

Finney, D.J. (1995): Thoughts suggested by a recent paper: Questions on non-parametric analysis of quantitative data (Letter to editor). J. Toxicol. Sci., **20**, 165-170.

Hamada, C., Yoshino, K., Matsumoto, K., Nomura, M. and Yoshimura, I. (1998): Three-type algorithm for statistical analysis in chronic toxicity studies. J. Toxicol. Sci., **23**, 173-181.

Inaba, T. (1994): Problem of multiple comparison tests for evaluation of enzyme inhibitor XI. Bull. Jap. Soc. Biopharm. Stat., **40**, 33-36.

Jonckheere, A. R. (1954): A distribution-tree k-sample test against ordered alternatives. Biometrika, **41**, 133-145.

Kobayashi, K., Watanabe, K. and Inoue, H. (1995): Questioning the usefulness of the nonparametric analysis of quantitative data by transformation into ranked data in toxicity studies. J.Toxicol. Sci., **20**, 47-53.

Kobayashi, K. (1997a): A comparison of one- and two-sided tests for judging significant differences in quantitative data obtained in toxicological bioassay of laboratory animals. J. Occup. Health, **39**, 29-35.

Kobayashi, K., Ohori, K., Kobayashi, M. and Takeuchi, H. (1997b): Choice of method for statistical analysis of quantitative data obtained from toxicological studies. San Ei Shi, **39**, 86-92 (in Japanese).

Kobayashi, K., Ohori, K., Kanamori, M. and Takeuchi, H. (1997c): Comparison to Dunnett and *t*-test for two groups. Jap. Soc. Occup. Health, Tokai Branch, pp.56-57 (No. 203).

Kobayashi, K., Miura, D., Ohori, K., Kobayashi, M., Kanamori, M. and Takeuchi, H. (1997d): Statistics for grade data of pathological findings. 43th. Tokai-koshueisei-gagai, (No. 3-03).

Kobayashi, K., Miura, D., Watari, N., Yamamoto, T., Shoji, A., Kitajima, S., Shiya, M., An, R., Sho, H. and Takasaki, K. (1998): A view of Bartlett test for small sized animal setting. Shizuoka Experimental Animals Research Association, **25**, 8-13 (in Japanese).

Kobayashi, K., Kanamori, M., Ohori, K. and Takeuchi, H. (2000a): A new decision tree method for statistical analysis of quantitative data obtained in toxicity studies on rodent. San Ei Shi, **42**, 125-129 (in Japanese).

Kobayashi, K. (2000b): Trends of the decision tree for selecting hypothesis-testing procedures for the quantitative data obtained in the toxicological bioassay of the rodents in Japan. J. Environ. Biol., **21**, 1-9.

NIH(2006): http://ntp.niehs.nih.gov/index.cfm?objectid=08480 1F0-F43F-7B74-0BE549908B5E5C1C

MHLW (2006): http://wwwdb.mhlw.go.jp/ginc/html/db1-j.html

Nagano, K., Umeda, Y., Senoh, H., Gotoh, K., Arito, H., Yamamoto, S. and Matsushima, T. (2006): Carcinogenicity and chronic toxicity in rats and mice exposed by inhalation to 1, 2-dichloroethane for two years. J. Occup. Health, **48**, 424-436.

OECD (1995): Organization for Economic Cooperation and Development. OECD Guideline for the Testing of Chemicals. 407 Adopted 27.07.95. Repeated Dose 28-day Oral Toxicity Study in Rodents. Geneva.

Sakaki, H., Igarashi, S., Ikeda, T., Imamizo, K., Omichi, T., Kadota, M., Kawaguchi, T., Takizawa, T., Tsukamoto, O., Terai, K., Tozuka, K., Hirata, J., Handa, J., Mizuma, H., Murakami, M., Yamada, M. and Yokouchi, H. (2000): Statistical method appropriate for general toxicological studies in rats. J. Toxicol. Sci., **25**, 71-98 (in Japanese).

Sano, M. and Okayama, Y. (1990): Programs of Dunnett's multiple comparison test by one and two sided with high accuracy table. Bull. Jap. Soc. Biopharm. Stat., **32**, 21-44 (in Japanese).

Sumida, K., Saito, K., Ueda, K., Otsuka, M., Tsujimura, K., Miyaura, H., Sekijima, M., Nakayama, K., Kawano, Y., Kawakami, Y., Asamoto, M. and Shirai, T. (2006): Optimization of an animal test protocol for toxicogenomics studies (II). J.Toxicol. Sci., **32**, 47-56.

USEPA (2000): United States Environmental Protection Agency. Health Effects Test Guidelines. OPPTS 870. 3050. Repeated Dose 28-Day Oral Toxicity Study in Rodents. Washington, D.C., USA.

USFDA (2000): United States Food and Drug Administration. Short-Term Toxicity Studies with Rodents. Toxicological Principle for the Safety Assessment of Food Ingredients. Redbook 2000. Rockville, USA.

Yamazaki, M., Noguchi, Y., Tanda, M. and Shintani, S. (1981): Statistical method appropriates for general toxicological studies in rats. J. Takeda Res. Lab., **40** (3/4), 163-187 (in Japanese).

Yoshimura, I. and Ohashi, S. (1992): Statistical Analysis for Toxicology Data. pp23, Chijin-shokan, Tokyo (in Japanese).

Yoshimura, I. (1987): In "Statistical analysis of toxicological data", pp45, Scientist, Tokyo (in Japanese).

Mutation Research 655 (2008) 1–3

Contents lists available at ScienceDirect

Mutation Research/Genetic Toxicology and Environmental Mutagenesis

journal homepage: www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Comparison of different methods for an accurate assessment of cytotoxicity in the *in vitro* micronucleus test
I. Theoretical aspects

E. Lorge [a,*], M. Hayashi [b], S. Albertini [c], D. Kirkland [d]

[a] Servier Group, Drug Safety Assessment, BP 43255, 45403 Fleury-les-Aubrais, France
[b] Division of Genetics and Mutagenesis, NIHS, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan
[c] F. Hoffmann-LaRoche AG, Department of Toxicology, CH-4070 Basel, Switzerland
[d] Covance Laboratories Limited, Otley Road, Harrogate HG3 1PY, UK

ARTICLE INFO

Article history:
Received 28 January 2008
Received in revised form 28 May 2008
Accepted 3 June 2008
Available online 14 June 2008

Keywords:
Cytotoxicity
Genotoxicity
In vitro micronucleus test
Mouse lymphoma cells
Population doubling
Relative increase in cell count
Increased cell count

ABSTRACT

A decrease in the cytokinesis-block proliferation index (CBPI) or replication index (RI) is routinely used to determine cytotoxicity of a test compound and therefore the choice of its appropriate test concentration for the *in vitro* micronucleus (MN) test conducted in the presence of cytochalasin B. As a number of laboratories prefer to conduct the *in vitro* MN test in the absence of cytochalasin B, it is important that selected test concentrations, based on cytotoxicity, should be similar to what they would have been if cytochalasin B had been used, and should be relevant of a true cytotoxicity. By using models to analyse the dynamics of the cell cultures with and without cytochalasin B we have compared different methods for evaluation of cytotoxicity, and demonstrate that relative decrease in population doubling or relative increase in cell counts are the most appropriate measures of cytotoxicity to compare with reduction in CBPI or RI.

1. Introduction

The overall cytotoxicity of an agent in a cell culture results in both cytostasis and cell death. Cytostasis is not restricted to inhibitors of cell division but may also be the consequence of many cytotoxicity pathways leading to a delayed cell cycle. As cytostasis may result from effects on cell division and may also be cell death-related, this component of cytotoxicity could be predominant and therefore should be accurately taken into account in the overall cytotoxicity.

In the *in vitro* micronucleus test, when using cytochalasin B, two main methods for measurement of cytotoxicity have been recommended [1,2]. One is based on CBPI (Cytokinesis-Block Proliferation Index), where:

$$\mathrm{CBPI} = \frac{\text{no. mononucleated cells} + 2 \times \text{no. binucleated cells} + 3 \times \text{no. multinucleated cells}}{\text{total number of cells}}$$

CBPI is determined in treated ($\mathrm{CBPI_T}$) and control ($\mathrm{CBPI_C}$) cultures, and the amount of cytostasis induced by the treatment is determined as follows:

$$\%\,\text{cytostasis} = 100 - 100\left[\frac{\mathrm{CBPI_T} - 1}{\mathrm{CBPI_C} - 1}\right].$$

Another recommended measure is the replication index (RI). This directly gives the extent of cell replication in treated cultures relative to control as follows:

$$\mathrm{RI} = \frac{(\text{no. binucleated cells} + 2 \times \text{no. multinucleated cells})/\text{total number of cells treated cultures}}{(\text{no. binucleated cells} + 2 \times \text{no. multinucleated cells})/\text{total number of cells control cultures}} \times 100$$

Thus, the percentage of cytostasis = 100 − RI.

Many laboratories prefer to conduct the *in vitro* micronucleus test in the absence of cytochalasin B, so the above measures of CBPI and RI cannot be used. To avoid irrelevant positive results due to excessive cytotoxicity or other artifacts [3], it is important that selection of the top concentration (based on inducing a certain level of toxicity) should be similar in the absence or presence of cytochalasin B.

* Corresponding author. Tel.: +33 2 38 23 86 11.
E-mail address: elisabeth.lorge@fr.netgrs.com (E. Lorge).

1383-5718/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrgentox.2008.06.003

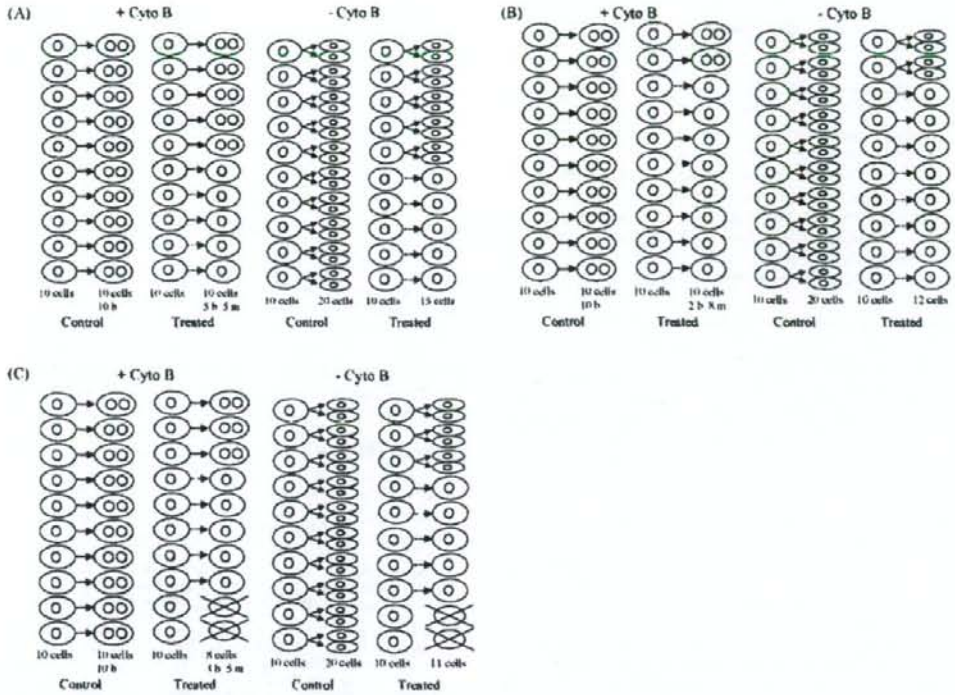


Fig. 1. Theoretical examples of cytotoxicity. (A) First example, inhibition of division in 50% of cells. (B) Second example, inhibition of division in 80% of cells (high concentration). (C) Third example, inhibition of cell division and cell death. b: binucleated; m: mononucleated.

Several different methods to measure cytotoxicity in the absence of cytochalasin B are available, but no clear recommendations have been made other than general statements in the IWGT report [1]. A similar effort of clarification has already been made in case cytochalasin B is used [2], but it is also required when cytochalasin B is not used. One commonly used method is the determination of relative cell counts (final counts in treated cultures relative to final counts in control cultures). However, this static measurement does not reflect all types of cytotoxicity, and therefore may lead to selection of excessively high concentrations, which may lead to irrelevant *in vitro* positive results [3]. Also, cytotoxicity may be different according to the method with which it is evaluated [4], so it is important to assess different measurements for their appropriateness. There are methods that can be used in the absence of cytochalasin B that better reflect the ability of the cultures to proliferate, for example relative population doubling (RPD):

$$\text{RPD} = \frac{\text{no. of population doublings in treated cultures}}{\text{no. of population doublings in control cultures}} \times 100,$$

Table 1
Comparison of different indicators of cell proliferation (and by deduction, indicators of cytotoxicity) in the presence or absence of cytochalasin B

	+CytoB		-CytoB		
	CBPI	RI	Population doubling	Increase in cell counts	Cell counts
	Cytokinesis-block proliferation index	Replicative index			
First example (Fig. 1A)					
Value in control	2.0	1.0	1.0	10.0	20.0
Value in treated	1.5	0.5	0.6	5.0	15.0
Relative cytostasis or cytotoxicity[a]	50.0%	50.0%	41.5%	50.0%	25.0%
Second example (Fig. 1B)					
Value in treated	1.2	0.2	0.3	2.0	12.0
Relative cytostasis or cytotoxicity[a]	80.0%	80.0%	73.7%	80.0%	40.0%
Third example (Fig. 1C)					
Value in treated	1.4	0.4	0.1	1.0	11.0
Relative cytostasis or cytotoxicity[a]	62.5%	62.5%	86.2%	90.0%	45.0%

[a] Calculated according to the formulae in the text.

where population doubling = [log(post-treatment cell number/initial cell number)]/log 2 or relative increase in cell counts (RICC):

$$\text{RICC} = \frac{\text{increase in number of cells in treated cultures (final - starting)}}{\text{increase in number of cells in control cultures (final - starting)}} \times 100$$

We decided to compare these two measures and the more traditional relative cell counts, with CBPI and RI in several different kinetic models. In the examples below, we show that kinetic evaluation of cell growth by RPD or RICC is preferable to simple cell counts and is the equivalent of CBPI-derived parameters used in the presence of cytochalasin B. This modeling exercise was paralleled by the publication of real data using different cytotoxicity measurements in the accompanying paper [5].

2. Modeling of cytotoxicity with and without cytochalasin B

In Fig. 1 we have drawn some theoretical examples of different types of cytotoxic effect in the presence and absence of cytochalasin B. In the first example, we simply study the situation where 50% of the cells have failed to divide. In the second example the effects of a higher concentration of the test substance, where 80% of the cells failed to divide, are examined. In the third example, in addition to inhibition of cell division, some cells have also died. Table 1 compares the CBPI-derived replication index in the presence of cytochalasin B with RPD, RICC and relative cell counts in the absence of cytochalasin B.

In all three examples, relative cell counts give a much higher estimate of survival (i.e. lower estimate of cytotoxicity) than RPD, RICC, CBPI or RI. In the second and third example, the use of relative cell counts would have led to the selection of too high concentrations, on the basis of 50% cytotoxicity, whereas RPD, RICC, CBPI and RI would all have resulted in selection of a lower concentration, since treatment with the concentration in the examples caused toxicity in the range 62.5–90%. In the first and second example, RPD and RICC (58.5 and 50%, respectively) are similar to both RI and CBPI (50%), whilst relative cell counts indicate a much higher toxicity (75%). In the third example, where cell death occurs, RPD and RICC derive from RI. This is due to the fact that, in the absence of cytochalasin B, RPD and RICC, do not distinguish individual cells that have divided. On the other hand, this apparent discrepancy is also due to the fact that RI is related to the total number of harvested cells, in this example decreased by two cells that died, whilst RPD and RICC are related to the initial number of treated cells, which remains constant. These are the limitations of both measurements and both methods.

These theoretical case studies show that relative cell counts are the least favourable cytotoxicity measurement to use in the *in vitro* micronucleus assay. This is not surprising as, in the number of cells counted on harvest, the loss of cells due to cytotoxicity is offset by the cell gain due to cell division occurring during treatment. In fact, at 50% reduction in cell counts too few cells have divided in the cultures to allow a meaningful evaluation of micronuclei in the absence of cytochalasin B.

In conclusion, simple cell counts were shown to underestimate cytotoxicity and, therefore, should not be recommended for the evaluation of cytotoxicity in the absence of cytochalasin B, as this method is likely to lead to the selection of excessively high top concentrations, susceptible to induce irrelevant positive results. Instead it is recommended that reduction in population doubling or reduction in relative increase in cell counts be used as cytotoxicity measures in the absence of cytochalasin B.

Conflicts of Interest

None.

References

[1] M. Kirsch-Volders, T. Sofuni, M. Aardema, S. Albertini, M. Fenech, M. Ishidate Jr., S. Kirchner, E. Lorge, T. Morita, H. Norppa, J. Surralles, A. Vanhauwaert, A. Wakata, Report from the in vitro micronucleus assay working group, Mutat. Res. 540 (2003) 153–163.

[2] M. Kirsch-Volders, T. Sofuni, M. Aardema, S. Albertini, D. Eastmond, M. Fenech, M. Ishidate Jr., S. Kirchner, E. Lorge, T. Morita, H. Norppa, J. Surralles, A. Vanhauwaert, A. Wakata, Corrigendum to "report from the in vitro micronucleus assay working group", Mutat. Res. 564 (2004) 97–100.

[3] D. Kirkland, M. Aardema, L. Henderson, L. Mueller, 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.

[4] M.D. Fellows, M.R. O'Donovan, Cytotoxicity in cultured mammalian cells is a function of the method used to estimate it, Mutagenesis 22 (2007) 275–280.

[5] M.D. Fellows, M.R. O'Donovan, E. Lorge, D. Kirkland, Comparison of different methods for an accurate assessment of cytotoxicity in the *in vitro* micronucleus test. II. Practical aspects with toxic agents, Mutat. Res. 655 (2008) 4–21.