Published in final edited form as:

Mutat Res. 2007 December 1; 634(1-2): 235–240.

Sensitivity of the Erythrocyte Micronucleus Assay: Dependence on Number of Cells Scored and Inter-animal Variability

Grace E. Kissling^{1,*}, Stephen Dertinger², Makoto Hayashi³, and James T. MacGregor⁴

¹National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709

²Litron Laboratories, Rochester, NY 14623

³National Institute of Health Sciences, Tokyo 158-8501, Japan

⁴Toxicology Consulting Services, Arnold, MD 21012

Abstract

Until recently, the *in vivo* erythrocyte micronucleus assay has been scored using microscopy. Because the frequency of micronucleated cells is typically low, cell counts are subject to substantial binomial counting error. Counting error, along with inter-animal variability, limit the sensitivity of this assay. Recently, flow cytometric methods have been developed for scoring micronucleated erythrocytes and these methods enable many more cells to be evaluated than is possible with microscopic scoring. Using typical spontaneous micronucleus frequencies reported in mice, rats and dogs, we calculate the counting error associated with the frequency of micronucleated reticulocytes as a function of the number reticulocytes scored. We compare this counting error with the inter-animal variability determined by flow cytometric scoring of sufficient numbers of cells to assure that the counting error is less than the inter-animal variability, and calculate the minimum increases in micronucleus frequency that can be detected as a function of the number of cells scored. The data show that current regulatory guidelines allow low power of the test when spontaneous frequencies are low (e.g., $\leq 0.1\%$). Tables and formulas are presented that provide the necessary numbers of cells that must be scored to meet the recommendation of the International Working Group on Genotoxicity Testing that sufficient cells be scored to reduce counting error to less than the inter-animal variability, thereby maintaining a more uniform power of detection of increased micronucleus frequencies within each species across the range of spontaneous frequencies observed in these three species.

Keywords

Erythrocyte micronucleus assay; flow cytometry; binomial counting error; inter-animal variability; power calculation

1. Introduction

The ability of the *in vivo* erythrocyte micronucleus assay to detect small increases in the spontaneous background frequency of micronucleated cells in a group of animals (or study subjects) is limited by either the binomial counting error, when the number of cells scored gives small numbers of scored events (micronucleated cells) [1–4], or by inter-animal variation, when that variation is so large that it obscures a small, but real, increase. Furthermore, the sensitivity

* Corresponding author: Dr. Grace E. Kissling, (919) 541-1756 (phone), (919) 541-4311(FAX), kissling@niehs.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

to detect small increases in the micronucleus frequency in an individual animal is limited at small micronucleus counts by the binomial counting error and at higher counts by the spontaneous variability in that individual animal over the time span in which measurements are made. Recognizing these facts, the working group on the *in vivo* micronucleus assay organized by the International Workshops on Genotoxicity Testing (IWGT) has recommended that, whenever possible, sufficient cells should be scored to reduce the counting error to less than the variability in MN frequency between individual animals (for comparison of values in different treated groups) [5].

Prior to the development of flow cytometric scoring methods, the number of cells scored was generally limited by the practical consideration of the number of cells that could be scored in a reasonable period of time by a microscopist, and therefore the minimum number of cells recommended to be scored in current regulatory guidelines is generally less than that required to discern differences between individual animals. Flow cytometric methodologies now make it practical to reduce the counting error to very small values [6–8], allowing, for the first time, reliable determination of the inter- and intra-animal variation in the spontaneous micronucleus frequency.

We summarize below experimentally determined mean and variability among animals in the spontaneous frequency of micronucleated reticulocytes (MN-RETs) in peripheral blood reticulocytes in the Sprague-Dawley rat, CD-1 mouse, and beagle dog, and in bone marrow reticulocytes in the Sprague-Dawley rat, and compare this inter-animal variability with the microscopic counting error associated with the current regulatory recommendations for scoring bone marrow or peripheral blood reticulocytes in these species. From these values, we determine the minimum increase in group mean frequencies of MN-RETs that can be detected in these species, tabulate the minimum increases that can be detected as a function of the number of RETs scored, and identify the numbers of cells that need to be scored to meet the IWGT recommendation that sufficient cells should be scored such that the error in individual animal MN-RET frequencies is less than the inter-animal variability.

2. Inter-Animal Variability

The inter-animal variability of the percentage of MN-RETs among RETs ($\#MN-RET/\#RETs \times 100$) in the peripheral blood of Sprague-Dawley rats, CD-1 mice, and purpose-bred beagle dogs, and also in the bone marrow of Sprague-Dawley rats after removal of nucleated cells on a cellulose mini-column as described by Romagna [9] and Weiner et al. [10], was estimated by scoring 20,000 reticulocytes using the flow cytometric method described by Dertinger et al. [11–13]. These data are summarized in Table 1. The data are taken from previously reported studies in these species [14,15]. Details of the experimental methodology are reported in the previous publications. As is discussed below, scoring 20,000 RETs results in a sufficient number of events (MN-RETs) that the error associated with individual animals does not exceed approximately 50% of the inter-animal variability of spontaneous MN-RET frequencies in the respective species. The inter-animal % coefficient of variation ($\%CV = \text{Standard Deviation} \div \text{Mean} \times 100\%$) of the MN-RET frequencies was 41% for the rat, 35% for the mouse, and 30% for the dog.

Table 2 presents the binomial error in the count of MN cells in an individual animal obtained by scoring 2000, 4000, 8000, or 20,000 RETs as a function of the spontaneous frequency of MN-RETs. It should be noted that the spontaneous frequency in rodent bone marrow or peripheral blood reported by different experienced laboratories has ranged from 0.05% in rat (see individual laboratory values in [14]) to a mean value of 0.2% in the mouse [15,16] and 0.31% in the beagle dog. Since the counting error depends on the background rate and the number of cells scored, we have tabulated values over the range of spontaneous frequencies

commonly reported in rodents and recently observed in the beagle dog (manuscript in preparation). As can be seen in Table 2, when 2000 cells are scored (the recommended number in the current OECD, FDA, and EPA regulatory guidelines [17–19] the error in the counts observed in individual animals is substantially greater than the variation between animals (Table 1). When the spontaneous frequency is 0.1%, approximately 6000 cells would need to be scored to reduce the error in the individual animal count to less than the inter-animal variability observed in the rat.

3. Sensitivity to Increases above the Spontaneous Frequency

Table 3 summarizes the minimum increases above the spontaneous frequency that can be detected in groups of five animals (the minimum currently recommended in OECD, FDA, and EPA guidelines [17–19] as a function of the number of target cells scored (in this case RETs) and observed spontaneous frequency (in this case %MN-RETs among RETs). Minimum detectable increases in MN-RET frequencies at $p \leq 0.05$ or $p \leq 0.01$, with 90% or 95% power were determined using Monte Carlo simulations. Specifically, to reflect inter-animal variability, five binomial probabilities were randomly selected from a normal distribution with the following mean, μ_0 , and standard deviation, σ , combinations: $(\mu_0, \sigma) = (0.05\%, 0.02\%)$, $(0.10\%, 0.045\%)$, $(0.20\%, 0.070\%)$, or $(0.30\%, 0.092\%)$. For a given fold-increase, f , a second set of five binomial probabilities were randomly selected from a normal distribution with mean, $\mu_1 = \mu_0 \times f$, and the same σ given above. Using the 5 binomial probabilities from the spontaneous mean group, five MN-RETs frequencies were randomly generated from binomial distributions, with n = number of RETs scored, 2000, 4000, or 20,000. Such selection from a binomial distribution introduces the binomial counting error. Five MN-RET frequencies were similarly generated using the 5 binomial probabilities from the increased mean group. A one-tailed Mann-Whitney test was then performed on these 10 counts, comparing the spontaneous group to the increased group, and the p-value was noted as to whether it was 0.05 or less and/or 0.01 or less. This was repeated 3000 times and the percentages of the 3000 ‘samples’ for which the p-value was 0.05 or less and 0.01 or less were calculated. The process was repeated over a series of increases, f , at increments of 0.1, to determine the first point at which the power exceeded 90% or 95%. We obtained very similar results (not shown) by generating the 5 binomial probabilities from beta distributions having the above combinations of μ_0 , μ_1 and σ .

For the line labeled “ ∞ ” in Table 3, there is no counting error; rather, the variability in frequencies is due to inter-animal variation alone. If we assume that inter-animal variation is normally distributed, the minimum difference between μ_1 and μ_0 , $\delta = \mu_1 - \mu_0$, detectable using 5 animals per group with significance level α and power $1 - \beta$ is

$$\delta = (t_{\alpha} + t_{\beta})\sigma\sqrt{\frac{2}{5}}$$

[20]. Here, t_{α} and t_{β} are the critical values from the $5 + 5 - 2 = 8$ degree of freedom t -distribution having upper tail probabilities α and β , respectively. The minimum detectable fold-increase over the spontaneous group is then

$$f = \frac{\mu_1}{\mu_0} = \frac{\mu_0 + \delta}{\mu_0} = 1 + \frac{\delta}{\mu_0}.$$

While spontaneous MN-RET frequencies determined from counting 2000 RETs from different animals are not often normally distributed, it has been our experience that spontaneous frequencies determined from counting 20,000 RETs from different animals are approximately normally distributed. Therefore, the assumptions of normality that we made above are most likely reasonable.

It should be noted that even if the counting error of the MN-RET frequency in each individual animal could be eliminated, the sensitivity of detection of changes in the observed mean group frequency would still be limited by the inter-animal variability (represented in Table 3 by the line in which an infinite number of cells is scored). It is clear that the regulatory assay as currently conducted is relatively insensitive to changes in the spontaneous frequency, especially when the spontaneous frequency is low. For example, when the spontaneous frequency is 0.05% and only 2000 RETs are scored, even a 6.8-fold increase would fail to be detected at a confidence level of $p \leq 0.01$ in 10% of experiments conducted. Even at the more commonly-reported spontaneous frequency of 0.1% a 4.8-fold increase would fail to be detected 10% of the time at this same confidence level. The use of flow cytometric scoring to achieve a sufficient cell count to allow individual animal frequencies with adequate certainty (*i.e.*, certainty of the individual value relative to the inter-animal variation) would increase the sensitivity such that a doubling of a spontaneous frequency of 0.1% among 20,000 RETs scored would be detected nearly 90% of the time at a confidence level of $p \leq 0.05$. It should also be noted that, regardless of the spontaneous frequency, the sensitivity achieved by scoring 20,000 RETs is close to the optimal sensitivity that could be achieved if no counting error were present.

4. Number of Reticulocytes Required to be Scored to Reduce Counting Error to Less than Inter-Animal Variability

Table 4 summarizes the number of cells required to be scored to reduce the counting error of individual animal values (Table 2, %CV) to the observed inter-animal variation or less (Table 1, %CV). These numbers were calculated by setting a multiple ($m = 1.0, 0.5, \text{ or } 0.2$) of the inter-animal %CV equal to the binomial counting error %CV and solving for the required sample size, n . Mathematically, if p is the percent of MN-RETs among all RETs within an animal, then

$$\% CV_{\text{binomial error}} = \frac{\sqrt{\frac{p(1-p)}{n}}}{p} = m \times \% CV_{\text{inter_animal}}$$

Solving for n , we get

$$n = \frac{1-p}{p(m \times \% CV_{\text{inter_animal}})^2}.$$

The numbers of RETs required are prohibitively laborious to obtain by conventional microscopic scoring, but are easily achieved by automated procedures such as flow cytometry.

Acknowledgements

We thank Ronald Fiedler of Pfizer, Inc. for providing the data on spontaneous frequencies and inter-animal variability of the frequency of micronucleated reticulocytes in the bone marrow of Sprague-Dawley rats. This research was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences.

References

1. Margolin BH, Collings BJ, Mason JM. Statistical Analysis and Sample-Size Determinations for Mutagenicity Experiments With Binomial Responses. *Environ Mutagen* 1983;5:705–716. [PubMed: 6617600]
2. Margolin, BH.; Risko, KJ. The statistical analysis of in vivo genotoxicity data: case studies of the rat hepatocyte UDS and mouse bone marrow micronucleus assays. In: Ashby, J.; de Serres, FJ.; Shelby, MD.; Margolin, BH.; Ishidate, M., Jr; Becking, GC., editors. *Evaluation of Short-term Tests for Carcinogens: Report of the International Programme on Chemical Safety's Collaborative Study on In*

- Vivo Assays. 1. Cambridge University Press on behalf of WHO; Cambridge and New York, NY: 1988. p. 1.29-1.42.
3. Kastenbaum MA, Bowman KO. Tables for determining the statistical significance of mutation frequencies. *Mutat Res* 1970;9:527–549. [PubMed: 5424720]
 4. Hayashi M, Yoshimura I, Sofuni T, Ishidate M Jr. A procedure for data analysis of the rodent micronucleus test involving a historical control. *Environ Mol Mutagen* 1989;13:347–359. [PubMed: 2737186]
 5. Hayashi M, MacGregor JT, Gatehouse DG, Blakey DH, Dertinger SD, Abramsson-Zetterberg L, Krishna G, Morita T, Russo A, Asano N, Suzuki H, Ohyama W, Gibson D. *In vivo* erythrocyte micronucleus assay: III. Validation and regulatory acceptance of automated scoring and the use of rat peripheral blood reticulocytes, with discussion of non-hematopoietic target cells and a single dose-level limit test. *Mutat Res* 2007;627:10–30. [PubMed: 17157053]
 6. Asano N, Torous D, Tometsko C, Dertinger S, Morita T, Hayashi M. Practical threshold for micronucleated reticulocyte induction for low doses of mitomycin C, Ara-C and colchicine. *Mutagenesis* 2006;21:15–20. [PubMed: 16364928]
 7. Grawe J, Abramsson-Zetterberg L, Zetterberg G. Low dose effects of chemicals as assessed by the flow cytometric *in vivo* micronucleus assay. *Mutat Res* 1998;405:199–208. [PubMed: 9748577]
 8. Torous D, Asano N, Tometsko C, Sugunan S, Dertinger S, Morita T. Performance of flow cytometric analysis for the micronucleus assay - a reconstruction model using serial dilutions of malaria-infected cells with normal mouse peripheral blood. *Mutagenesis* 2006;21:11–13. [PubMed: 16188876]
 9. Romagna F. Current issues in mutagenesis and carcinogenesis; fractionation of a pure PE and NE population from rodent bone marrow. *Mutat Res* 1988;206:307–309. [PubMed: 3200254]
 10. Weiner SK, Fiedler RD, Schuler MJ. Development and Evaluation of a Flow Cytometric Method for the Analysis of Micronuclei in Rat Bone Marrow *In Vivo*. *Environ Molec Mutagen* 2004;43:236.
 11. Dertinger SD, Torous DK, Tometsko K. Simple and reliable enumeration of micronucleated reticulocytes with a single-laser flow cytometer. *Mutat Res* 1996;371:283–292. [PubMed: 9008730]
 12. Dertinger SD, Torous DK, Hall NE, Tometsko CR, Gasiewicz TA. Malaria-infected erythrocytes serve as biological standards to ensure reliable and consistent scoring of micronucleated erythrocytes by flow cytometry. *Mutat Res* 2000;464:195–200. [PubMed: 10648906]
 13. Dertinger SD, Camphausen K, MacGregor JT, Bishop ME, Torous DK, Avlasevich S, Cairns S, Tometsko CR, Menard C, Muanza T, Chen Y, Miller RK, Cederbrant K, Sandelin K, Pontén I, Bolcsfoldi G. Three-color labeling method for flow cytometric measurement of cytogenetic damage in rodent and human blood. *Environ Mol Mutagen* 2004;44:427–35. [PubMed: 15517570]
 14. MacGregor JT, Bishop ME, McNamee JP, Hayashi M, Asano N, Wakata A, Nakajima M, Aidoo A, Moore MM, Dertinger SD. Flow cytometric analysis of micronuclei in peripheral blood reticulocytes:II. An efficient method of monitoring chromosomal damage in the rat. *Toxicol Sci* 2006;94:92–107. [PubMed: 16888079]
 15. Torous DK, Hall NE, Illi-Love AH, Diehl MS, Cederbrant K, Sandelin K, Pontén I, Bolcsfoldi G, Ferguson LR, Pearson A, Majeska JB, Tarca JP, Hynes GM, Lynch AM, McNamee JP, Bellier PV, Parenteau M, Blakey D, Bayley J, van der Leede BM, Vanparys P, Harbach PR, Zhao S, Filipunas AL, Johnson CW, Tometsko CR, Dertinger SD. Interlaboratory validation of a CD71-based flow cytometric method (MicroFlow®) for the scoring of micronucleated reticulocytes in mouse peripheral blood. *Environ Mol Mutagen* 2005;45:44–55. [PubMed: 15605355]
 16. Heddle JA, Salamone MF, Hite M, Kirkhart B, Mavournin K, MacGregor JT, Newell GW. The induction of micronuclei as a measure of genotoxicity. *Mutat Res* 1983;123:61–118. [PubMed: 6888413]
 17. OECD, Guideline for the testing of chemicals. Mammalian erythrocyte micronucleus test. Guideline 474, July, 1997.
 18. FDA (U.S. Food and Drug Administration). Office of Food Additive Safety, Redbook 2000, Toxicological principles for safety assessment of food ingredients. Updated November 2003, www.cfsan.fda.gov/~redbook/red-toca.html
 19. EPA (U.S. Environmental Protection Agency) Health Effects Test Guidelines OPPTS 870.5395. Mammalian Erythrocyte Micronucleus Test, Office of Prevention, Pesticides and Toxic Substances

- (7101) EPA 712-C-98-226. Aug. 1998 www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-5395.pdf
20. Snedecor, GW.; Cochran, WG. Statistical Methods. 6. The Iowa State University Press; Ames, IA: 1976. p. 113-116.
 21. Fiedler, R. Pfizer, Inc, Personal communication. 2007.
 22. Beutler E, Gelbart T. The mechanism of removal of leukocytes by cellulose columns. Blood Cells 1986;12:57-64. [PubMed: 3790738]

Mean and Inter-Animal Variation of the Micronucleated Reticulocyte Frequency in the Peripheral Blood (PB) and Bone Marrow (BM) of Sprague-Dawley Rats, Swiss Mice, and Beagle Dogs.

Table 1

Species	Strain/ Breed	Tissue	Mean %MN- RET	Std. Dev. of % MN-RET	Inter-animal % CV ^b	# Animals	# Expts.	Reference
Rat	SD	PB	0.11	0.045	41%	15	3	[14]
Rat	SD	BM	0.23 ^a	0.059 ^a	26%	190	38	[21] R. Fiedler; Personal communication
Mouse	CD-1	PB	0.20	0.070	35%	79	9	[15]
Dog	Beagle	PB	0.31	0.092	30%	22	4	Manuscript In Preparation

^a Bone marrow %MN-RET values were determined by separation of nucleated cells on a mini-cellulose column [9,22], with subsequent scoring of the MN-RET frequency among 20,000 RETs by the same flow cytometric procedure used for analysis of peripheral blood.

^b %CV = Std. Dev./Mean ×100%

Table 2
Counting Error (Standard Deviation (S.D.) and Coefficient of Variation) of Individual Animal Values of MN-RET Frequency as a Function of True Spontaneous Frequency and Number of RETs Scored.

True %MN-RET	# RETs Scored per Animal	S.D. of Count	Counting Error %CV
0.05	2000	0.050	100%
	4000	0.035	71%
	8000	0.025	50%
	20,000	0.016	32%
0.10	2000	0.071	71%
	4000	0.050	50%
	8000	0.035	35%
	20,000	0.022	22%
0.20	2000	0.100	50%
	4000	0.071	35%
	8000	0.050	25%
	20,000	0.032	16%
0.30	2000	0.122	41%
	4000	0.086	29%
	8000	0.061	20%
	20,000	0.039	13%

Minimum Detectable Increases in MN-RET Frequency in Groups of Five Animals as a Function of Spontaneous Frequency and Number of RETs Scored^a

Spontaneous Frequency (%MN-RET)	Number of RETs Scored	Minimum Detectable Fold Increase in Spontaneous Frequency		
		With 90% Probability at p ≤ 0.05	With 95% Probability at p ≤ 0.05	at p ≤ 0.01
0.05 S.D. = 0.020	2000	4.5	6.8	9.3
	4000	3.5	5.5	6.3
	20,000	2.3	3.1	3.4
0.10 S.D. = 0.045 (rat PB)	∞	1.8	2.1	2.2
	2000	3.3	4.8	6.4
	4000	2.9	4.2	4.7
0.20 S.D. = 0.070 (mouse PB)	20,000	2.2	3.0	3.2
	∞	1.8	2.0	2.4
	2000	2.7	3.9	2.1
0.20 σ = 0.059 (rat BM)	4000	2.3	3.2	4.5
	20,000	1.9	2.5	3.5
	∞	1.7	2.0	2.7
0.30 S.D. = 0.092 (dog PB)	2000	2.7	3.9	2.1
	4000	2.2	3.1	4.4
	20,000	1.8	2.3	3.3
	∞	1.6	1.8	2.5
	2000	2.4	3.4	1.9
	4000	2.1	2.8	3.7
	20,000	1.8	2.3	3.0
	∞	1.6	1.8	2.4
				1.9

^aValues for the cases of infinite cell counts are calculated based on the observed inter-animal variability (standard deviation from Table 1) for the species stated, assuming no counting error; the inter-animal variability for frequency 0.05% is assumed to be 0.02%. The detectable increase depends on the relative magnitudes of both the counting error and the inter-animal variability. Although counting error can be reduced by scoring more RETs, the minimum detectable increase cannot go below a bound determined by the inter-animal variability (i.e., the value given in the infinite cell count rows). Species entries correspond to the approximate spontaneous frequency and associated inter-animal standard deviation in the species specified in Table 1.

Number of Reticulocytes Required to be Scored to Reduce Counting Error to Less Than the Observed Inter-Animal Coefficient of Variation.

Table 4

Spontaneous Frequency (% MN- RET)	Species/Tissue With this Appx. Spontaneous Freq,	Inter-Animal % CV ^a	Number of RETs to be scored to reduce counting error %CV to:	
			Equal the inter-animal % CV	50% of the inter-animal % CV
0.05	Rat BM & PB(Microscopy, some reports)	NA ^b	NA	NA
0.10	Rat PB (Data cited above)	41%	5943	23,772
0.20	Mouse BM & PB	35%	4074	16,294
0.30	Dog	30%	3693	14,770
				148,573
				101,837
				92,315

^a Experimentally-determined inter-animal %CV by flow cytometric scoring of 20,000 peripheral blood RETs, at the approximate spontaneous frequency tabulated.

^b Inter-animal % CV has not been determined at the spontaneous frequency of 0.05%; no reported experiments have scored sufficient cells to determine the inter-animal variability.