Available online at www.sciencedirect.com
SCIENCE DIRECT®
Mutation Research 534 (2003) 45–64
Genetic Toxicology and Environmental Mutagenesis
www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Intra- and inter-laboratory variation in the scoring of micronuclei and nucleoplasmic bridges in binucleated human lymphocytes Results of an international slide-scoring exercise by the HUMN project

Michael Fenech [a], Stefano Bonassi [b,*], Julie Turner [a], Cecilia Lando [b], Marcello Ceppi [b], Wushou Peter Chang [c], Nina Holland [d], Micheline Kirsch-Volders [e], Errol Zeiger [f], Maria Paola Bigatti [g], Claudia Bolognesi [h], Jia Cao [i], Giuseppe De Luca [j], Marina Di Giorgio [k], Lynnette R. Ferguson [l], Aleksandra Fucic [m], Omar Garcia Lima [n], Valeria V. Hadjidekova [o], Patrizia Hrelia [p], Alicja Jaworska [q], Gordana Joksic [r], A.P. Krishnaja [s], Tung-Kwang Lee [t], Antonietta Martelli [u], Michael J. McKay [v], Lucia Migliore [w], Ekaterina Mirkova [x], Wolfgang-Ulrich Müller [y], Youichi Odagiri [z], Thierry Orsiere [a1], Maria Rosaria Scarfì [a2], Maria J. Silva [a3], Toshio Sofuni [a4], Jordi Suralles [a5], Giorgio Trenta [a6], Irena Vorobtsova [a7], Anne Vral [a8], Andrea Zijno [a9]

[a] *CSIRO Health Sciences and Nutrition, P.O. Box 10041, Adelaide, SA 5000, Australia*
[b] *Department of Environmental Epidemiology, Istituto Nazionale per la Ricerca sul Cancro, L. go R. Benzi 10, I-16132 Genova, Italy*
[c] *Institute of Environmental Health Sciences, National Yang Ming University Medical School, 155 Section 2 Lih-non Street, 11200 Taipei, Taiwan*
[d] *School of Public Health, University of California, 217 Warren, Berkeley, CA 94720-7360, USA*
[e] *Laboratory for Cell Genetics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium*
[f] *Errol Zeiger Consulting, Chapel Hill, NC 27517, USA*
[g] *Dipartimento di Biologia Animale e dell' Uomo, University of Turin, V. Accademia Albertina 17, 10123 Turin, Italy*
[h] *Toxicological Evaluation Unit, Istituto Nazionale per la Ricerca sul Cancro, L. go R. Benzi 10, 16132 Genova, Italy*
[i] *Molecular Toxicology Laboratory, Third Military Medical University, Chong Qing 400038, China*
[j] *ANPA, Via V. Brancati 48, 00144 Rome, Italy*
[k] *Laboratorio Dosimetria Biologica, Autoridad Regulatoria Nuclear, Av. del Libertador 8250, 1429 Capital Federal, Buenos Aires, Argentina*
[l] *Auckland Cancer Society Research Center, Private Bag 92019, Auckland 1000, New Zealand*
[m] *Institute for Medical Research end Occupational Health, 10000 Zagreb, Ksaverska C2, Croatia*
[n] *Centro de Proteccion e Higiene de las Radiaciones, Calle 20 No. 4109 e/41y47 Miramar, AP 6195 Habana 6, Cuba*
[o] *Laboratory Radiation Mutagenesis, National Center Radiobiology and Radiation Protection, Bul. Kl. Ochridski 132, Sofia 1756, Bulgaria*
[p] *Department of Pharmacology, University of Bologna, Via Irnerio 48, 40126 Bologna, Italy*
[q] *Norwegian Radiation Protection Authority, P.O. Box 55, 1332 Oesteraas, Norway*
[r] *Vinca Institute of Nuclear Sciences, Medical Protection Center, 300-110, P.O. Box 522, 11001 Belgrade, Yugoslavia*
[s] *Cell Biology Division, Bhabha Atomic Research Center, Mumbai 400 085, India*
[t] *Department of Radiation Oncology, Brody school of Medicine at ECU, Greenville, NC 27858, USA*
[u] *DIMI, V. le Benedetto XV 2, 16132 Genoa, Italy*

* Corresponding author.
E-mail address: stefano.bonassi@istge.it (S. Bonassi).

PII: S1383-5718(02)00248-6

[v] Divisions of Radiation Oncology and Research, Peter McCallum Cancer Institute, Smorgon Family Building, St. Andrews Place East Melbourne, Melbourne, Vic. 3002, Australia
[w] Dipartimento di Scienze dell'Uomo e dell'Ambiente, University of Pisa, Via S. Giuseppe 22, 56126 Pisa, Italy
[x] National Center of Hygiene, Medical Ecology and Nutrition, 15 Boul. D. Nestorov, Sofia 1431, Bulgaria
[y] Institute fur Medizinische Strahlenbiologie, Universitatsklinikum Essen, Hufelandstrasse 55, D45122 Essen, Germany
[z] Division of Human and Health Sciences, Yamanashi Prefectural College of Nursing, 1-6-1 Ikeda, Kofu 400-0062, Japan
[a1] Laboratoire de Biogenotoxicologie et de Mutagenese Env., Faculte de Medicine, Universite de la Mediterranee, 27 Bld. Jean Moulin, F-13385 Marseille Cedex 05, France
[a2] CNR-IRECE, Via Diocleziano 328, 80124 Naples, Italy
[a3] Centro de Genetica Humana, Portugal National Institute of Health, Av. Padre Cruz, 1649-016 Lisboa, Portugal
[a4] Chromosome Research Center, Olympus Optical Co. Ltd., 2-3 Kuboyama-cho, Hachioji-shi, Tokyo 192-8512, Japan
[a5] Dpt de Genetica i de Microbiologia, Facultat de Ciencies, Universitat Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[a6] DOSIME Group, Occupational Medicine Service, National Institute of Nuclear Physic, Via E. Fermi 40, 00044 Frascati (Rome), Italy
[a7] Central Research Institute of Roentgenology and Radiology, 189646 Leningradskaya Str. 70/4, Pesochny-2, St. Petersburg, Russia
[a8] Department of Radiation Protection, Faculty of Medicine, Proeftuinstraat 86, B-9000 Ghent, Belgium
[a9] Laboratory of Comparative Toxicology and Ecotoxicology, ISS, V. le Regina Elena 229, 00161 Rome, Italy

Received 6 May 2002; received in revised form 29 August 2002; accepted 20 September 2002

Abstract

One of the objectives of the HUman MicroNucleus (HUMN) project is to identify the methodological variables that have an important impact on micronucleus (MN) or micronucleated (MNed) cell frequencies measured in human lymphocytes using the cytokinesis-block micronucleus assay. In a previous study we had shown that the scoring criteria used were likely to be an important variable. To determine the extent of residual variation when laboratories scored cells from the same cultures using the same set of standard scoring criteria, an inter-laboratory slide-scoring exercise was performed among 34 laboratories from 21 countries with a total of 51 slide scorers involved. The results of this study show that even under these optimized conditions there is a great variation in the MN frequency or MNed cell frequency obtained by individual laboratories and scorers. All laboratories ranked correctly the MNed cell frequency in cells from cultures that were unirradiated, or exposed to 1 or 2 Gy of gamma rays. The study also estimated that the intra-scorer median coefficient of variation for duplicate MNed cell frequency scores is 29% for unexposed cultures and 14 and 11% for cells exposed to 1 and 2 Gy, respectively. These values can be used as a standard for quality or acceptability of data in future studies. Using a Poisson regression model it was estimated that radiation dose explained 67% of the variance, while staining method, cell sample, laboratory, and covariance explained 0.6, 0.3, 6.5, and 25.6% of the variance, respectively, leaving only 3.1% of the variance unexplained. As part of this exercise, nucleoplasmic bridges were also estimated by the laboratories; however, inexperience in the use of this biomarker of chromosome rearrangement was reflected in the much greater heterogeneity in the data and the unexplained variation estimated by the Poisson model. The results of these studies indicate clearly that even after standardizing culture and scoring conditions it will be necessary to calibrate scorers and laboratories if MN, MNed cell and nucleoplasmic bridge frequencies are to be reliably compared among laboratories and among populations.

Keywords: Micronucleus; HUMN project; Nucleoplasmic bridges; Scoring; Human lymphocytes; DNA damage; Variability; Protocol

1. Introduction

The HUman MicroNucleus (HUMN) Project is an international collaborative project which has three main goals: (a) compilation and comparison of base-line micronucleus (MN) frequencies in human populations to establish "normal" base-line frequencies of DNA damage and determine the main demographic, environmental, and methodological variables that impact on this index; (b) comparison of the various methods used to measure MN frequencies in human blood and epithelial cells to identify

Abbreviations: BN, binucleated; CBMN assay, cytokinesis-block micronucleus assay; MN, micronucleus or micronuclei; MONO, mononucleated; MULT, multinucleated; NDI, nuclear division index; NPB, nucleoplasmic bridge

important methodological variables and establish standard protocols to enable more reliable comparison of data among laboratories and among populations; and (c) to establish prospective epidemiological studies aimed at determining whether the MN frequency predicts risk of cancer and other degenerative diseases associated with DNA damage and ageing [1,2].

The HUMN Project was established in 1997 [1] and currently involves the participation of more than 35 laboratories worldwide. The impetus for this project was derived from the widespread use of the cytokinesis-block micronucleus (CBMN) assay to study DNA damage in human populations, and the recognition that by 1997 more than 12,000 measurements in human subjects had been reported in the literature using this method. Micronuclei are one of the best established biomarkers of chromosome damage. They are used in the in vitro testing of chemicals and radiation for genotoxicity [3–5] and also as an in vivo biomarker of exposure to genotoxins, and deficiency in folate and Vitamin B12 [6–13]. Micronuclei originate from chromosome fragments or whole chromosomes that lag at anaphase because they lack a centromere, or the centromere is defective, or there is a defect in the mechanism that enables the chromosomes to distribute correctly to the poles of the cell at anaphase [3–5]. As a consequence, MN are expressed in dividing cells. In human lymphocytes some MN are present as a result of a prior cell division in vivo, but most MN are expressed in lymphocytes after the cells are stimulated to divide in vitro. The CBMN assay relies on the observation that cells that have completed nuclear division and have had their cytokinesis blocked with cytochalasin, express chromosome damage as MN in the resulting binucleated (BN) cells [14–16]. It has recently been shown that it is possible to also score nucleoplasmic bridges (NPB), in addition to MN, in BN cells [3,5]. NPB are an important biomarker of chromosome rearrangement that is otherwise not measured in the CBMN assay, and originate from dicentric chromosomes that are pulled to opposite poles of the cell at anaphase [3,5].

The first objective of the HUMN project, i.e. compilation and comparison of data bases of base-line MN frequencies in human populations in cytokinesis-blocked lymphocytes, has already been completed and described [2]. This study of MN frequency in 6583 human subjects (measured in 25 laboratories distributed in 16 countries) verified that age and gender were the most important host factors influencing baseline MN frequency. A model based on the negative binomial function of probability was used to estimate the proportion of variability of MN frequency that can be explained by the studied parameters. The total amount of variability explained by the complete model (which included host factors, exposure to genotoxic agents, methodological parameters relating to culture, slide preparation and scoring criteria) was approximately 75%. Due to the correlation among parameters investigated, the sum of variance explained by all individual sources is greater than this value.

Because the CBMN assay is currently based on visual scoring of slides, the inter-scorer variation in the interpretation of scoring criteria and in the recognition of BN cells, MN, and NPB could account for a large proportion of the observed variability in the assay. Therefore, we designed an interlaboratory collaborative study to determine the extent of variability in the frequency of MN and other abovementioned endpoints as a function of visual scoring of slides by different laboratories and scorers, staining method and radiation dose.

The specific aims of this study were

1. To determine the extent of inter-laboratory variation in the MN assay scores when a common set of scoring criteria is used to score cells sampled from the same culture.
2. To determine the extent of intra-laboratory and intra-scorer variation in the CBMN assay.
3. To determine the effect of different staining methods on the micronucleus assay scores.
4. To evaluate the contribution of various parameters measured in the study to the total variability of the assay.

Other parallel aims were (a) to establish a comprehensive set of scoring criteria with an extensive set of photomicrographs to assist in the interpretation of slides (reported in accompanying paper [17]), and (b) to determine an acceptable coefficient of variation for duplicate measurements by a single scorer as a means for quality control of acceptable data in future studies.

2. Materials and methods

2.1. Study design

The study was advertised via the HUMN project email network. A total number of 34 laboratories from 21 countries indicated an interest in participating and each of these completed the study. Details about the laboratories, including geographical location, affiliation, principal investigators, number of scorers, and the staining method used is in Table 1.

The study design is shown schematically in Fig. 1. A blood sample was collected from a healthy 30-year-old male. Isolated lymphocyte cultures for the CBMN assay were prepared using published procedures [3,6,16]. Briefly, three cultures in RPMI 1640 with 10% foetal calf serum were prepared with a cell concentration of 1×10^6 ml^{-1}. Two of the cultures were exposed to gamma-rays (1 or 2 Gy at a dose rate of 5 Gy/min). The lymphocytes were stimulated to divide with phytohaemagglutin and 44 h later cytochalasin-B (4.5 ug/ml) (Sigma) was added and the cells harvested following an additional 28 h. A cytocentrifuge (Shandon) was used to prepare slides with two spots of cells on each slide. The cells were air-dried for 10 min and then fixed in absolute methanol for 10 min. Half of the slides were stained with Diff-Quik (Lab-Aids, Australia) and the rest left unstained and stored desiccated at 4 °C. The stained slides were mounted in DEPEX using a coverslip. Forty-seven stained and 47 unstained slides were prepared from each culture. The slides were coded using random numbers.

During September 2000, each laboratory received (by express courier) a package that contained (a) a detailed set of instructions, (b) one set of fixed and stained slides (control, 1 and 2 Gy), (c) one set of slides that were fixed but not stained (control, 1 and 2 Gy), (d) a detailed scoring procedure with diagrams and photomicrographs, and (e) a standard Excel template for recording MN and NPB scores. Any broken slides were replaced immediately with slides that had been made from the same culture. The unstained slides were stained upon arrival at the laboratories or kept desiccated at 4 °C until stained. Some laboratories ($N = 6$) had to alter their standard staining procedure by increasing staining time to adequately stain the fixed, but unstained, slides. A second batch of fixed but unstained slides were sent (to four laboratories) which were subsequently successfully stained in two laboratories but not in the other two, which were therefore unable to provide scores for these

Table 1
Laboratories that contributed data to the HUMN inter-laboratory method comparison (affiliation can be obtained from the authors list)

Country	Principal researcher(s)	Scorers	Staining method
Argentina	M. Di Giorgio	1	Giemsa
Australia	M. Fenech	4	Diff-Quik
Australia	M. McKay	1	DAPI[a]
Belgium	M. Kirsch-Volders	2	Giemsa
Belgium	H. Thierens and A. Vral	1	Azure B—Eosin
Bulgaria	V. Hadjidekova	1	Giemsa
Bulgaria	E. Mirkova	1	Giemsa
Croatia	A. Fucic	1	Giemsa
Cuba	O. Garcia Lima	1	Giemsa
France	T. Orsiere and A. Botta	1	Giemsa
Germany	W.U. Muller	2	Giemsa
India	A.P. Krishnaja	2	Giemsa
Italy	P.M. Bigatti	2	Giemsa[b]
Italy	C. Bolognesi	2	Giemsa
Italy	G. De Luca	1	Giemsa
Italy	P. Hrelia	1	Giemsa
Italy	A. Martelli	2	May-Grunwald, Giemsa
Italy	L. Migliore	2	Giemsa
Italy	M.R. Scarfì	2	Giemsa
Italy	G. Trenta	1	Giemsa
Italy	A. Zijno	1	Giemsa
Japan	Y. Odagiri	1	Giemsa
Japan	T. Sofuni	1	Acridine Orange[a]
New Zealand	L.R. Ferguson	1	Giemsa
Norway	A. Jaworska and K. Bredholt	2	Giemsa
Portugal	M.J. Silva	1	Giemsa
PR China	J. Cao Jia	3	Giemsa
Russia	I. Vorobtsova	1	May Grunwald, Giemsa
Spain	R. Marcos and J. Suralles	1	Giemsa
Taiwan	W.P. Chang	1	Hemacolor
USA	N. Holland	4	DAPI[a]
USA	T.K. Lee	1	Diff-Quik
Yugoslavia	G. Joksic	1	Giemsa

The order of the laboratories in this table does not reflect the assigned laboratory numbers.

[a] Indicates fluorescent staining method.

[b] Data not available for lab-stained slides because of difficulty with staining of the unstained slides provided.

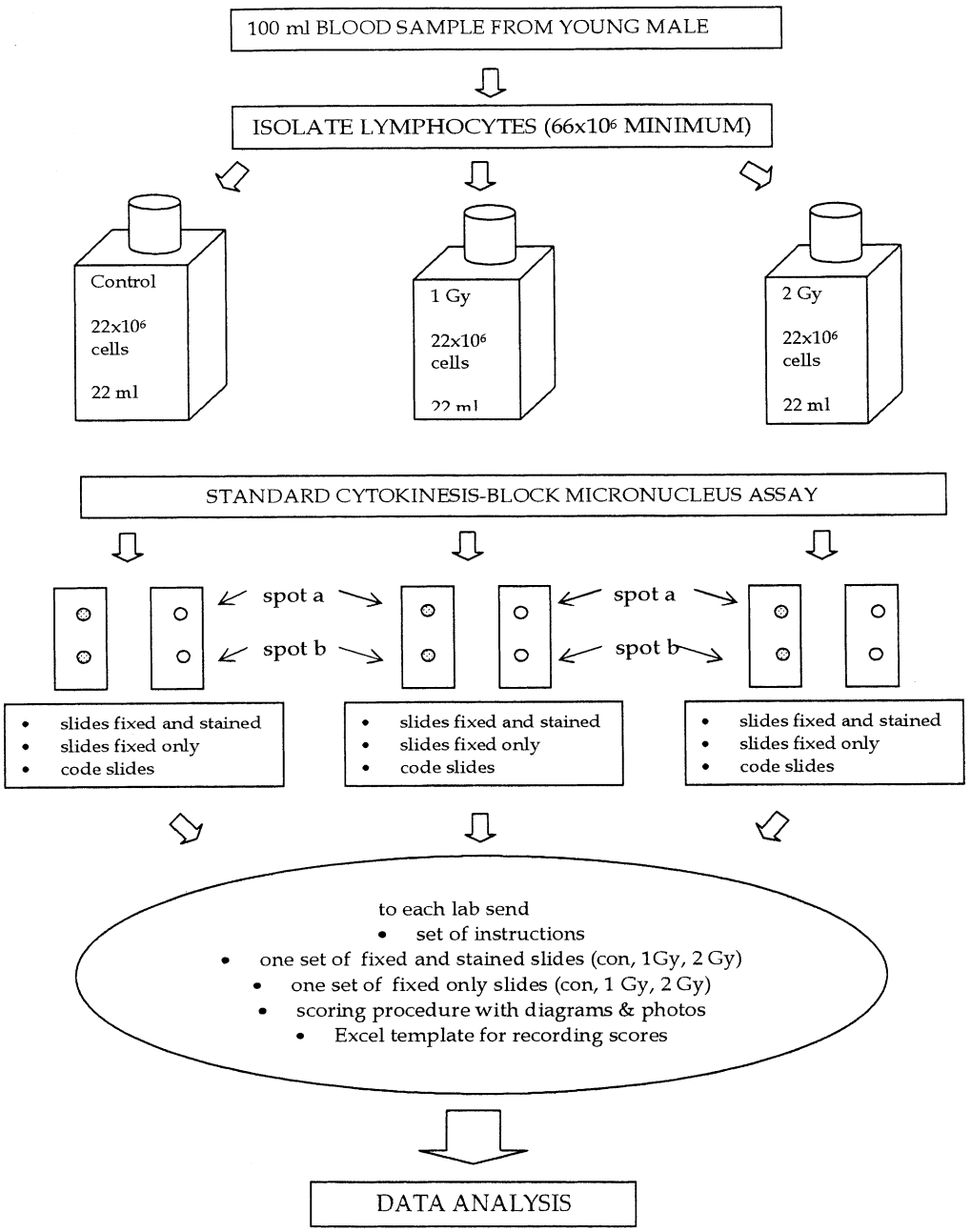


Fig. 1. Schematic diagram of the study design.

slides. Participants were advised to read the scoring procedure and scoring criteria carefully and requested not to proceed with scoring if they were uncertain about any aspect of the scoring procedure. If the laboratories had uncertainties regarding the procedures to be followed, they were directed to consult with Dr. Michael Fenech, who was responsible for preparation of the slides. Participants were advised to record the

data in the Excel template provided and to send the results to Dr. Stefano Bonassi by email by November 30, 2000 for analysis. The results from each laboratory were randomly allocated a code number from 1 to 34 to allow a blind statistical analysis and to allow descriptions of results from single laboratories. After data collection we were informed that one laboratory did not follow the scoring criteria provided; their scores were excluded from the study, which is therefore based on results from 33 laboratories.

2.2. Scoring procedure

Each laboratory could choose to have more than one person score the same set of slides, but a separate data sheet had to be generated for each scorer. Each scorer scored the full set of six slides (pre-stained and those stained in their laboratory). Each scorer obtained separate scores for the cells in the spot closest to the label on the slide (spot *a*) and the cells in the spot farthest from the label (spot *b*) (a total of 12 spots). It was requested that cells in spot *a* and spot *b* be scored on separate days to enable an analysis of the day to day variation in scoring and a derivation of the coefficient of variation for duplicate scores by one scorer. Information about the individual experience of scorers involved in the study was collected in terms of years of scoring and number of slides scored per year. This allowed the calculation of a new variable, i.e. number of slides scored in their lifetime, which was used in the following statistical analyses to measure the influence of a scorer's past experience.

For each spot of cells, the number of mononucleated (MONO) cells, binucleated (BN) cells, and multinucleated (MULT) cells in 500 viable cells were scored to determine the nuclear division index (NDI), as previously described [3,18]. It is important to note that participants were instructed not to score necrotic or apoptotic cells when determining the proportion of MONO, BN and MULT cells. They were also not required to distinguish between tri-nucleated and quadri-nucleated cells when scoring multinucleated cells because often it is hard to distinguish between the two and this adds an unnecessary burden to the assay.

For each spot of cells participants scored the following to determine chromosome damage:

1. The number of cells in 1000 BN cells containing one or more micronuclei.
2. The total number of micronuclei in 1000 BN cells.
3. The number of cells in 1000 BN cells containing one or more nucleoplasmic bridges (NPBs).

A BN cell that has both a micronucleus and an NPB was scored separately as a BN cell with a micronucleus and as a BN cell with a NPB.

Slides were examined at 1000× magnification for both light or fluorescence microscopy. Participants were requested not to score cells that they were unable to classify according to the criteria provided. They were asked to simply skip any cell that was unclassifiable and move on to the next. Similarly, MN and NPB where to be scored only if the scorer was confident that they met the given criteria.

2.3. Scoring criteria

A standard set of detailed scoring criteria, together with a comprehensive set of photomicrographs exemplifying the various types of cells, micronuclei, and nucleoplasmic bridges, were provided to each laboratory. These scoring criteria and the set of photomicrographs are described in the accompanying paper [17] and were based to a large extent on previously published procedures which however, did not provide photomicrographs [3,18].

2.4. Statistical methods

A Poisson regression analysis [19] was performed to estimate the contribution of each covariate to the total variance. According to the study design, the data have been considered as clusters which are hierarchically nested as follows: laboratory, scorer, staining, irradiation, and spot. Since observations *within* a cluster are (generally) more similar than observations *between* clusters, this lack of independence between data will lead to overdispersion, i.e. data showed more residual variability than expected according to an ordinary regression model [19,20]. Ignoring the clustered nature of data generally causes an underestimation of the variability of regression coefficients. A useful methodology that takes the hierarchical structure of the data into account is the multilevel approach [21]. The simplest multilevel model contains a

two-level structure, i.e. individuals (first level) within groups (second level). The model is

$$y_{ij} = \alpha + \alpha x_{ij} + \omega_j + \alpha_{ij}$$

where an observation y_{ij} referring to the ith individual in the jth group is modeled by the following parameters: α, the overall mean; β, the regression coefficient of the predictor variable x_{ij}; ω_j, the departure of the jth group mean from the overall mean; and ε_{ij}, the residual. In this model, α and β represent the fixed part of the model while ω_j and ε_j are the random part. In particular, we assume that ω_j follows a normal distribution with zero mean and variance σ^2_w. The multilevel approach consists in estimating α, β and σ^2_w. This last parameter is the index of heterogeneity among groups. This is a parsimonious method to quantify the differences among groups avoiding the estimate of a regression coefficient for each group. It is possible to evaluate the relative variance component for each level of clustering, i.e. the within-laboratory variability in the estimate of the parameter, the intra-laboratory-inter-scorer variability, the proportion of variance due to staining, and the heterogeneity between spots. Since our observations are counts, we can fit a Poisson multilevel model to our data. All analyses were carried out with GLIM [22] and MLwiN [23] statistical software.

3. Results

All participating labs completed the scoring of the pre-stained slides. Only 2 labs (#1 and 22) were unable to adequately stain the unstained slides provided and were therefore unable to provide data for this part of the study.

3.1. Descriptive statistics

A summary of results is reported in Table 2. Measures of central tendency (median and mean) and variability (S.D. and range) by dose of radiation are reported for each endpoint studied. Data from slides stained by the referent laboratory (pre-stained) and data from those stained by the preferred method of participating laboratories (lab-stained) were analyzed separately. An evident and significant positive association ($P < 0.01$) with the level of irradiation is present for all the genotoxicity end-points evaluated apart from NDI, which decreased with dose ($P < 0.01$). The mean frequency of MN per 1000 binucleated cells at 1 Gy is 8–11% higher that of micronucleated cells (MNed) at the background level. This difference is more pronounced at higher doses of irradiation, up to 34% at 2 Gy. The frequency of MN scored on pre-stained slides was generally higher

Table 2
Descriptive statistics concerning endpoints evaluated in the standardization study

End-point[a]	Dose Gray	Pre-stained slides (51 scorers)				Lab-stained slides (44 scorers)			
		Median	Mean	S.D.	Min–max	Median	Mean	S.D.	Min–max
MNed	0	8.0	8.8	4.3	2–28	7.0	7.4	3.3	1–19
	1	116.5	118.4	29.8	59–217	95.0	98.8	33.5	35–200
	2	305.9	302.2	69.9	165–563	257.5	255.4	77.6	63–564
MN	0	9.0	9.8	4.9	2–28	7.6	8.0	3.7	1–23
	1	138.5	140.8	37.1	61–265	105.0	115.3	39.5	36–230
	2	401.4	405.5	104.8	191–753	345.5	334.2	109.6	78–763
NPB	0	2.0	5.2	13.5	0–125	2.0	5.0	15.0	0–121
	1	18.0	22.3	18.1	3–124	16.8	19.6	16.0	0–93
	2	45.1	51.8	36.7	9–300	36.2	40.8	29.0	1–191
NDI	0	1.95	1.94	0.18	1.56–2.47	1.92	1.92	0.19	1.49–2.39
	1	1.89	1.88	0.19	1.44–2.37	1.86	1.84	0.19	1.43–2.21
	2	1.77	1.75	0.15	1.45–2.19	1.73	1.74	0.15	1.39–2.12

[a] MNed, micronucleated cells per 1000 BN cells; MN, micronuclei per 1000 BN cells; NPB, nucleoplasmic bridges per 1000 BN cells; NDI, nuclear division index.

than those of lab-stained slides, with a difference in mean values of 14% ($P < 0.01$).

Also in Table 2, it is evident that there is substantial overlap between MNed and MN in describing the effect of irradiation. Scoring the number of MN per cell gives a slightly better performance over scoring the frequency of MNed (with a regression coefficient, $\beta = 1.570$ versus 1.452 in the trend test by dose), although the improvement in precision is offset by a higher variability of the MN index (standard error (S.E.) 0.032 versus 0.028). Considering that most laboratories prefer to report the frequency of MNed cells, and given the need of selecting among the overwhelming amount of descriptive material we have produced, in the following parts of this paper we will describe in detail only the MNed cell frequency data, with the understanding that graphs and statistics presented correspond closely to those obtained for MN.

The scoring of NPB was an important aspect of the protocol. This endpoint is not widely used as yet, and all but one of the participating laboratories were inexperienced, having scored this biomarker for the first time within the framework of this study. A good association of NPB frequency with the irradiation dose is present ($\beta = 1.007$; $P < 0.01$), but the large extent of variability observed when compared with scoring MNed cells and MN cells for a more intense effort to standardize scoring procedures for this endpoint.

A broad picture of the results obtained from scoring MNed cells is given by Figs. 2A–F, which show the frequency of this parameter in the 33 laboratories involved in the analysis, by level of irradiation and slide preparation method. To provide information about the intra-laboratory variability, results of single scorers are reported for those laboratories with more than one scorer. As a measure of central tendency we chose the median, which is less affected by very high values, which are not uncommon in this kind of data. Laboratories are ranked according to the median value, which is represented by a closed black circle, with open circles representing single scorers. To provide direct information about the variability of values, lines corresponding to the overall median and 25th and 75th percentiles are reported in each figure.

The background values in non-irradiated cells (Fig. 2A and B), suggest that the large majority of laboratories are included between the lines corresponding to the critical percentiles. Values from the scoring of pre-stained slides are slightly higher, not only as median values, but also as extreme measures, with three laboratories reporting more than 17‰ MNed cells versus none in the set of laboratory-stained slides. Some laboratories tend to consistently score a higher frequency of MNed, i.e. laboratories 10, 24, 31, whereas other laboratories change their ranks dramatically according to the set of slides scored. This is the case for laboratories 25 and 34, with scores in the lower tertile for pre-stained slides and in the mid-high tertile when lab-stained slides are scored. The results from labs 18 and 29 are remarkable; they each obtained MNed frequencies that were two- and four-fold higher, respectively, in the pre-stained slides. In this latter case, however, extreme frequencies of MNed cells (and also of MN and NPB) are concentrated in one of the spots of cells on the slide, lending support to the hypothesis of a distribution anomaly when the cell suspension was spotted on the slides.

These inconsistencies in MNed cell frequency scores are also observed for slides of irradiated cultures, with some laboratories constantly at the right tail of the distribution, e.g. #5, 15, and other laboratories showing great changes of rank when scores from pre- and lab-stained slides are compared, e.g. #18, 24, 29, 31 (Fig. 2A–F).

The distribution of NPB scores is rather heterogeneous, with large differences among laboratories. Some laboratories, e.g. #24, 26, showed a greater variability with the presence of extreme values. Laboratory 20 had values generally 10–20-fold higher than the overall median. The repeated presence of extreme values is confirmed by the important difference between mean and median and the large dichotomy of min–max values in Table 2.

3.2. Inter- and intra-scorer variability

Fig. 2A–F also show the placement of each scorer in relation to their laboratory medians. Although a more analytic evaluation of inter-scorer variability within each laboratory is described later using multivariable hierarchical models, these figures provide interesting hints. Looking at the single combination of staining method and dose of irradiation, some heterogeneity becomes evident. A major aspect used to classify scorer's ability is the number of slides scored in their lifetime. In Table 3, the ratio of the median

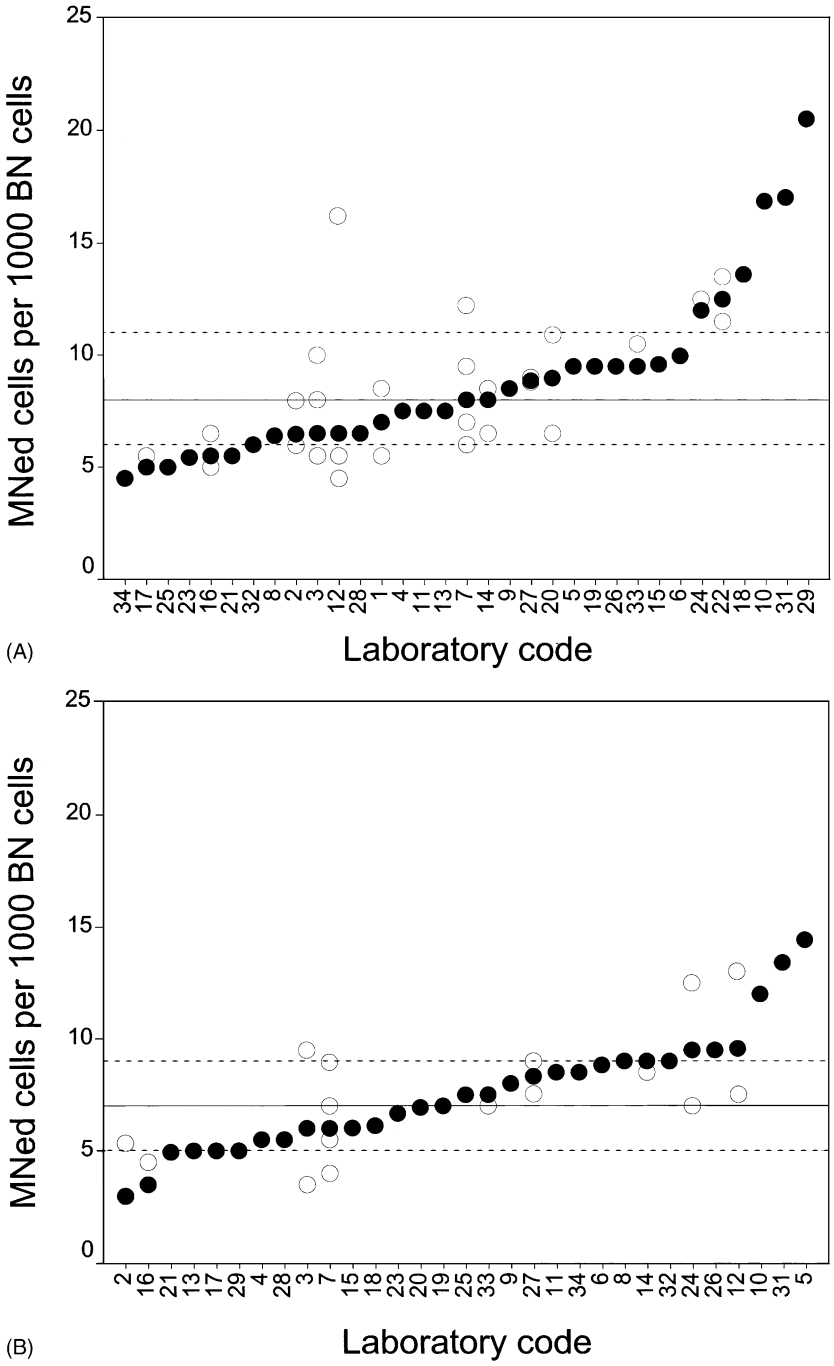
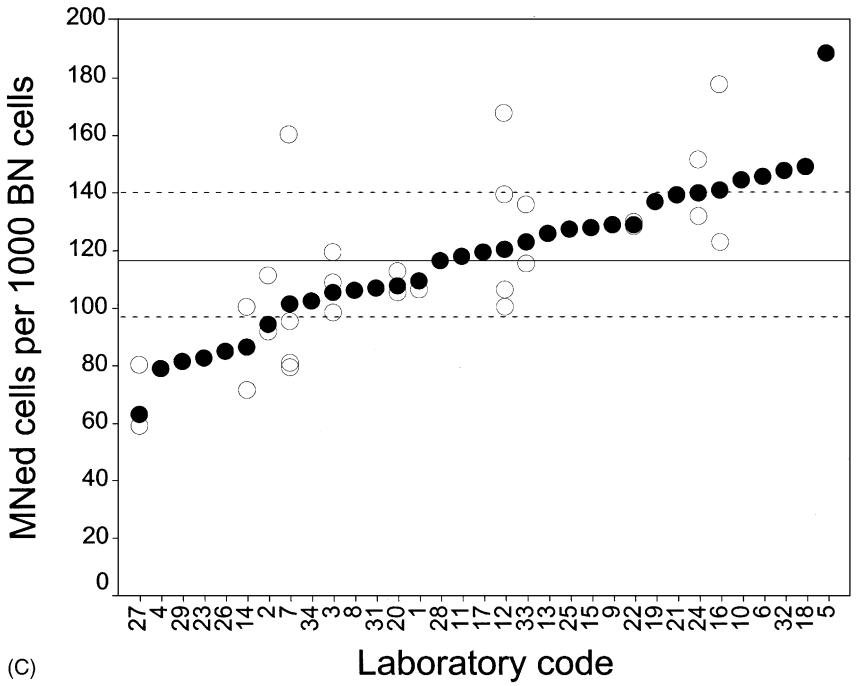
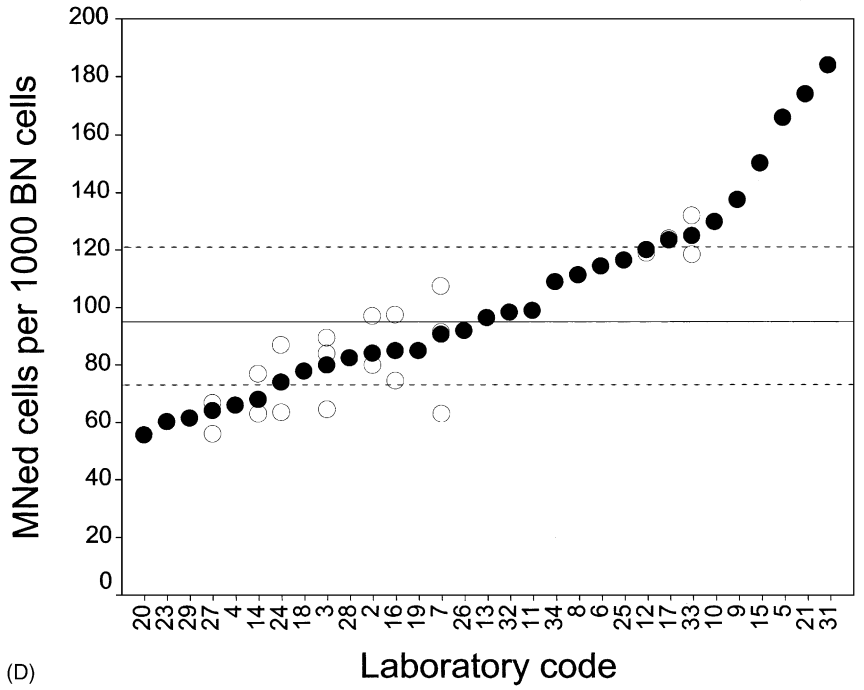


Fig. 2. (A) Laboratory MNed cell frequency scores (*y*-axis) for pre-stained slides from untreated cultures; (B) Laboratory MNed cell frequency scores (*y*-axis) for lab-stained slides from control cultures; (C) Laboratory MNed cell frequency scores (*y*-axis) for pre-stained slides from cultures exposed to 1 Gy; (D) Laboratory MNed cell frequency scores (*y*-axis) for lab-stained slides from cultures exposed to 1 Gy; (E) Laboratory MNed cell frequency scores (*y*-axis) for pre-stained slides from cultures exposed to 2 Gy; (F) Laboratory MNed cell frequency scores (*y*-axis) for lab-stained slides from cultures exposed to 2 Gy. The solid circles indicate the median of the results obtained by each laboratory that is identified by the number on the *x*-axis. The open circles indicate the results of individual scorers in laboratories that had more than one scorer completing the exercise. The solid line represents the overall median and the broken lines represent the 25th and 75th percentiles.

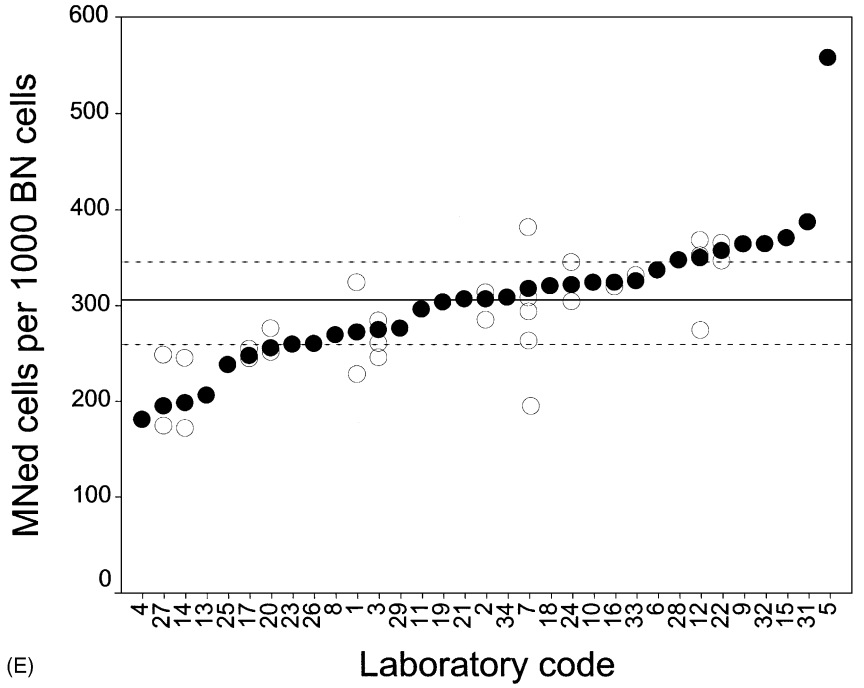


(C)

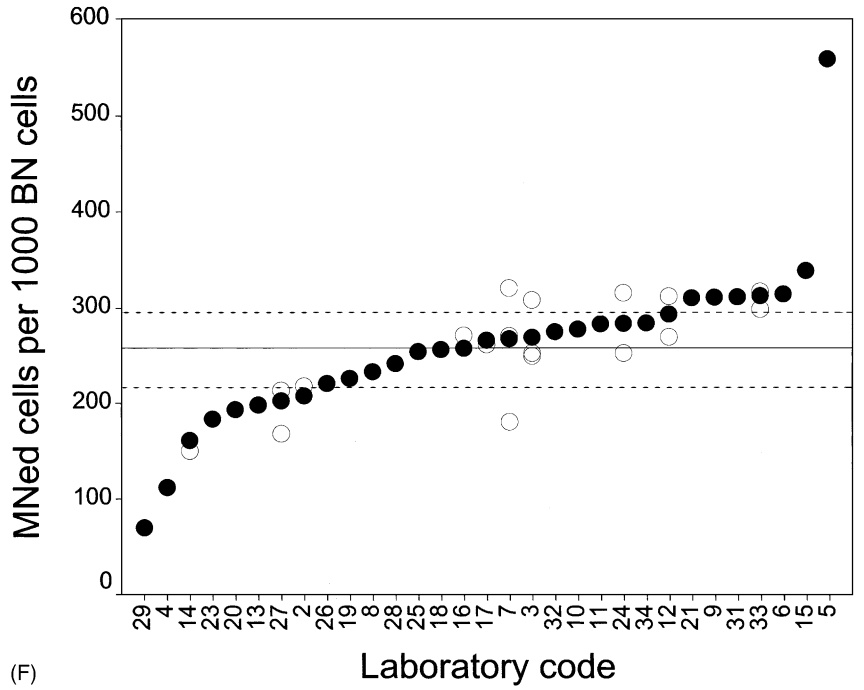


(D)

Fig. 2. (*Continued*).



(E)



(F)

Fig. 2. (*Continued*).

Table 3
The ratio of the median value of MNed cell frequency for 1 and 2 Gy cultures vs. the control culture by number of slides scored lifetime

Slides scored lifetime	Scorers (slides)	Median ratio		
		0 Gy	1 Gy	2 Gy
<51	4 (48)	1[a]	15.35	36.67
51–500	14 (143)	1.36	16.74	38.59
>500	21 (252)	1.24	17.71	44.70

[a] Reference level.

value of MNed cell frequency for 1 and 2 Gy cultures versus the control culture clearly reflects the number of slides scored previously, experienced scorers had a better capability to discriminate the effect of irradiation. Those laboratories with two scorers ($n = 10$) showed overall inter-scorer differences below 25%, with the most homogeneous scoring in laboratories 16 and 17. Laboratories with three or more scorers ($n = 4$) obviously produce more heterogeneity. A useful and straightforward measure to express the internal consistency of a scorer is the CV obtained from the discrepancy between spot *a* and spot *b* scores. This statistic, which is based on the ratio of the standard deviation to the mean, simply describes the variability as a proportion of the mean. This value turned out to be strictly dependent on the radiation dose, with an increasing precision of the estimates with increasing dose of exposure. Fig. 3 shows this dependency, revealing a real difference between irradiated cells, which have a median intra-scorer CV's for MNed of 14 and 11% at 1 and 2 Gy, respectively, versus the non-treated cells with a CV of 29%. A greater heterogeneity for the set of untreated slides is evident in Fig. 3, although some laboratories, e.g. laboratory #29, also had a great variability in irradiated cells.

These figures are consistent with those for MN. In contrast the scorer CV for NPB are much higher, with median scorer values of 82, 26, and 23% at 0, 1, and 2 Gy, respectively.

3.3. Measures of effect

In order to evaluate whether the MNed biomarker is concordant in evaluating the effect of radiation, despite the heterogeneity found among absolute measures of MNed frequency, the median MNed frequencies at 2 and 1 Gy were compared with the frequencies in the unexposed cells obtained from each laboratory. This exercise showed that irradiated cells experienced a higher frequency of MNed cells in pre- and lab-stained slides, in all laboratories, with a 14.3-fold higher frequency of MNed cells in cultures irradiated with 1 Gy, and an increase of 37.3-fold in those irradiated with 2 Gy. This pattern is overlapping in pre- and lab-stained slides for MNed cells and for MN. The distribution of single laboratories around these values is exemplified in Fig. 4. This figure refers to the median ratio values of MNed cell frequency for cells exposed to 1 or 2 Gy (relative to the control culture) based on all the slides scored (pre-stained and lab-stained) by each laboratory. All laboratories found an increasing MNed cell frequency by dose, although to different extents. The effect for 1 Gy ranged from 7.4- to 29.2-fold increase, and the 50% most frequent estimates were between 10.7 and 16.6, whereas exposing cells to 2 Gy caused a fold-increase of MNed cells ranging from 16.0 to 60.7, with the 50% most frequent estimates between 28.3 and 46.4.

The evaluation of NPB frequency in lab-stained slides describes non-homogeneously the effect of irradiation. In this set of slides, three laboraratories scored more NPB in the 1 Gy slides than in the 2 Gy slides, and two laboratories scored the same number of NPB at the two doses. This was not the case with the pre-stained slides, where the differential effect of two doses was always correctly recognized. In general, however, there was a good correlation between MNed and NPB frequency in both pre- (0.63; $P < 0.01$) and lab-stained slides (0.57; $P < 0.01$). Six laboratories reported that radiation effects were similar for NPB and MNed cells; this suggests that, with appropriate training, the sensitivity and reliability of NPB scoring can improve.

3.4. NDI

No evident differences in NDI were found between pre- and lab-stained slides. On the other hand, NDI was negatively associated with the dose of irradiation ($\beta = -0.090$, $P < 0.01$), decreasing from a mean of 1.93 in the untreated cells, to 1.86 at 1 Gy and 1.75 in the 2 Gy set. The overall mean of this index was 1.85, with a S.D. of 0.19; values ranged between 1.39 and 2.47 (see Table 2).

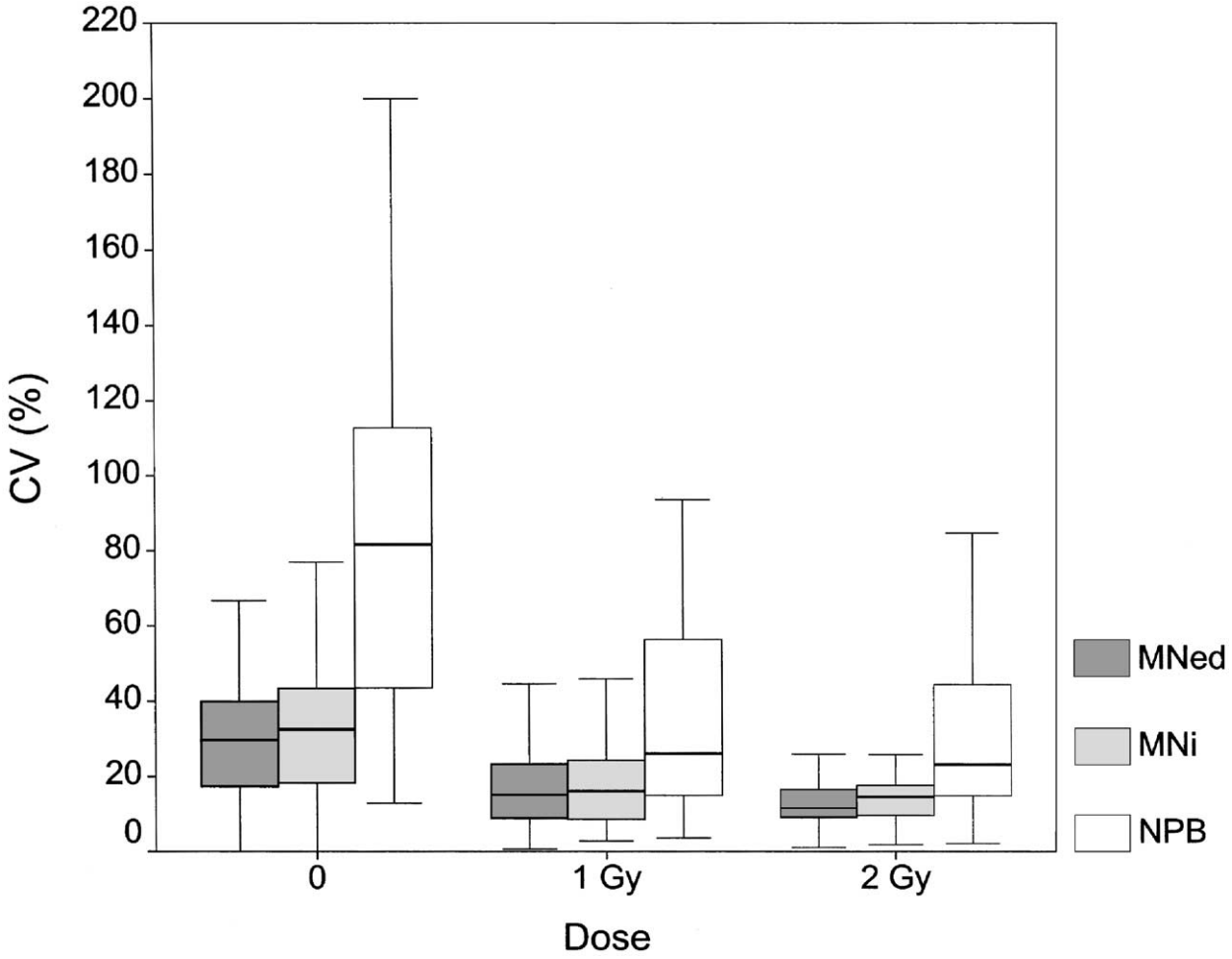


Fig. 3. Intra-scorer coefficient of variation for spot *a* and spot *b*. MNed, MN, and NPB frequency results for slides from unexposed, 1 and 2 Gy cultures. The results shown as box and whisker plots represent the median, the 25th and 75th percentile, and range for data of all scorers ($N = 51$).

3.5. Proportion of explained variability

One of the most important aspects in the standardization of an assay is the quantitative definition of sources of variability. To provide an evaluation of the contribution of various parameters to the general variability of data, a Poisson regression model that included the main effects of all studied variables was fitted to the whole database. A backward extraction of variables from this model allowed an estimation of the contribution of each variable in terms of variability. The results of this procedure are reported in Figs. 5 and 6. The percentages reflect the independent contribution of each variable, but because these parameters are not orthogonal (i.e. not independent), a portion of variability is explained by the correlation of two or more variables in the model. This part has been described as covariance.

From Fig. 5, it is evident that there is a very low level of unexplained variance for MNed frequency, i.e. 3.1%, partly due to the experimental design, which took the main explanatory variables into consideration,

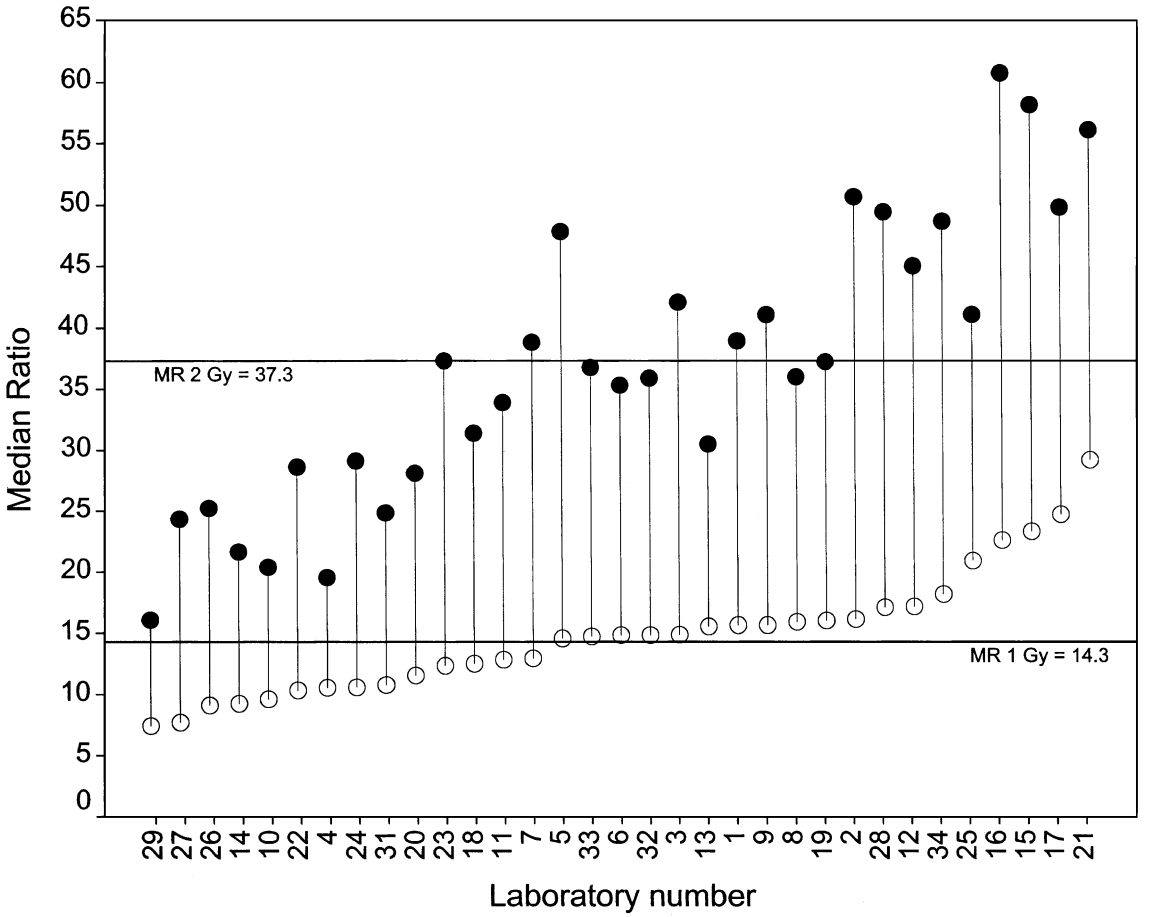


Fig. 4. Open and solid circles represent the median ratio values for cells exposed to 1 and 2 Gy, respectively, vs. the control culture shown for each laboratory, based on all the slides (pre- and lab-stained) scored. The horizontal lines represent the mean ratio values for 1 and 2 Gy. The numbers on the *x*-axis are the laboratory code number. Note: the vertical lines between the open and closed circles are of no mathematical significance and are simply used to help distinguish results of a single laboratory.

but also to the good agreement among participating laboratories. The main response variable, i.e. the dose of irradiation, is responsible for almost two-thirds of the explained variability. This is an important observation supporting the sensitivity of the test for the monitoring of this kind of exposure. The contribution of laboratory to the variance is quite small, and this is not surprising given that the individual contributions of other variables responsible for the differences among laboratories, such as NDI, scorer, spot, and staining method, are independently considered. The remaining variance is due to external and unknown factors. More than a quarter of the variability is due to the covariance between variables.

A different pattern is evident for NPB, where the proportion of unexplained variance is over 20%. The contribution of the dose is only one-third, and the 28.4% is attributable to unknown inter-laboratory factors (Fig. 6). These results confirm the need of improving training and scoring procedures for this endpoint.

The role of different methods used for laboratory staining of the slides has been evaluated by simply comparing the fluorescent versus non-fluorescent methods. A 15% lower frequency of MNed cells (95% CI; 6–23%) was found in the slides stained with fluorescent methods. For the three laboratories that chose fluorescence microscopy only 27% of the laboratory mean scores were within the 25th and 75th

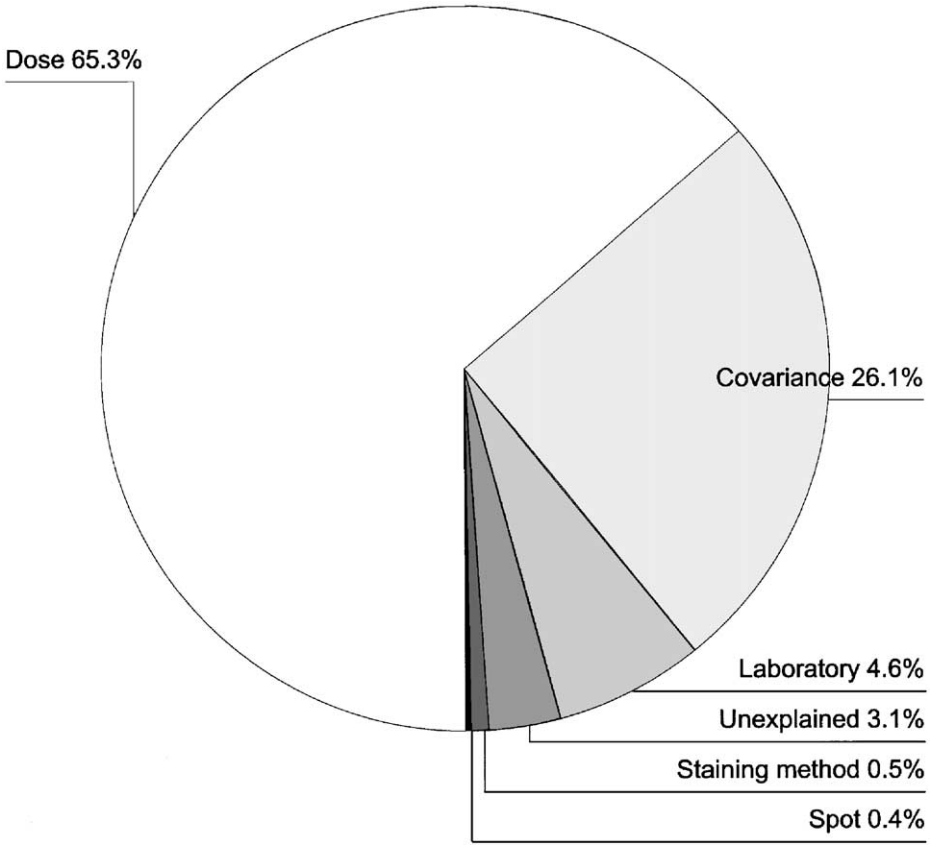


Fig. 5. Proportion of variance explained by variables in the model describing MNed frequency.

percentile compared to 56% for the laboratory mean scores from laboratories that chose light microscopy methods for scoring.

3.6. Measures of error

The design of this experiment has produced a typical hierarchical structure of data, with each level of aggregation nested within the following level. In each laboratory, scorers read slides stained by different methods, slides contained cells treated at different doses of radiation, and each slide had two cell samples, each on a separate spot. At each level of clustering it is possible to estimate variance components, e.g. if we consider inter-laboratory variability, it is possible to estimate how much of this variability is due to the scorer, to slide treatment, to the irradiation dose, and to sampling error between the two spots, and how much is due to external factors. This analysis has been performed in a subset of the whole database, i.e. in those laboratories with more than one scorer, to permit the evaluation of the variability due to this parameter. Variance components estimates have been calculated using random effects models, and the result of this analysis is shown in Table 4. If we look at the MNed assay we see that all components investigated, excluding the scorer, showed a remarkable degree of heterogeneity. The presence of this remaining uncertainty, despite the inclusion in the model of main potential predictors, means that there are other variables responsible for inter-laboratory differences that have not been considered in our model. On the other hand, if we look at the variable describing the scorer we see that this parameter is not significant. The reason is that all variability due to the scorer is completely explained by differences in staining methods, dose,

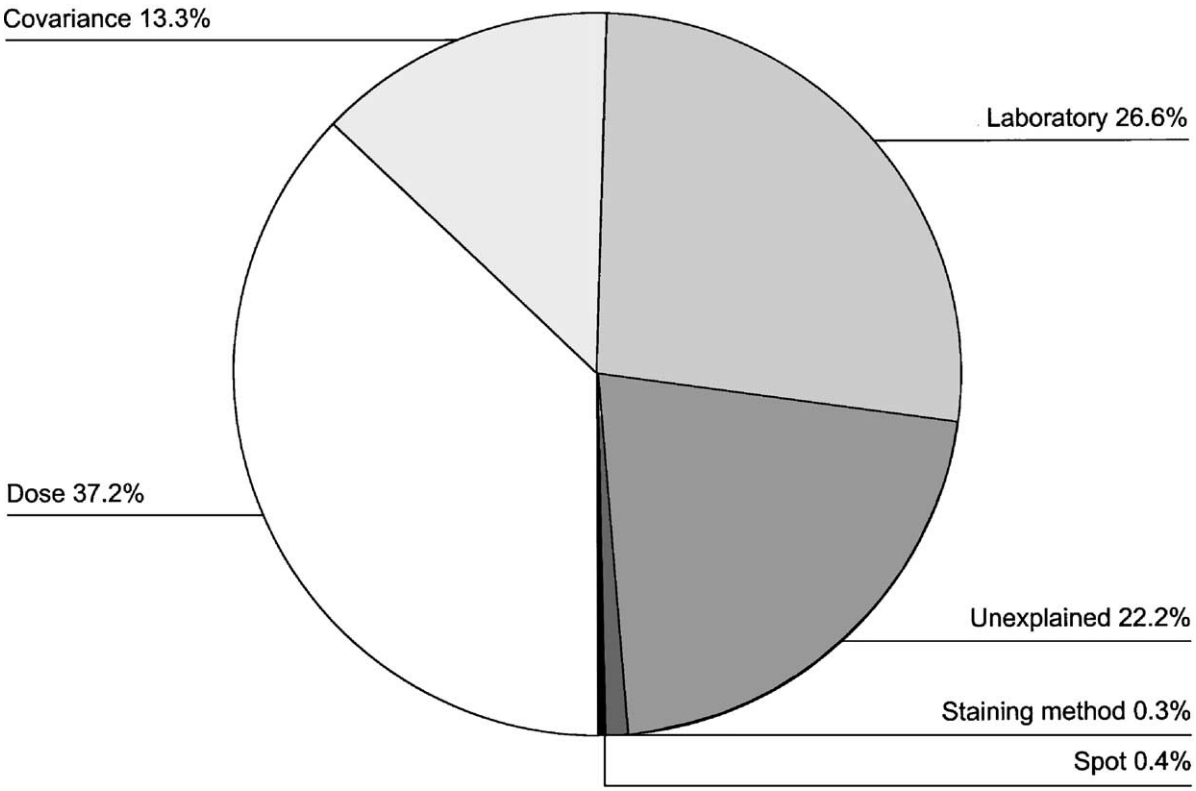


Fig. 6. Proportion of variance explained by variables in the model describing NPB frequency.

Table 4
Fixed and random coefficients of the multilevel hierarchical model for the three end-points considered

Model	MNed	MN	NPB
Fixed part			
Log median ratio (S.E.)			
Intercept	−4.766 (0.181)	−4.578 (0.199)	−3.641 (0.891)
Dose_1	2.594 (0.044)[a]	2.641 (0.044)[a]	1.220 (0.183)[a]
Dose_2	3.599 (0.045)[a]	3.721 (0.046)[a]	1.921 (0.195)[a]
NDI	−0.053 (0.089)	−0.093 (0.098)	−0.683 (0.434)
Random part			
Variance component (S.E.)			
Laboratory	0.023 (0.011)[a]	0.026 (0.014)	0.982 (0.441)[a]
Scorer	0.001 (0.006)	0.0001 (0.007)	0.0001 (0.0001)
Stain	0.016 (0.007)[a]	0.019 (0.009)[a]	0.210 (0.117)
Dose	0.012 (0.004)[a]	0.014 (0.004)[a]	0.663 (0.121)[a]
Spot	0.008 (0.002)[a]	0.014 (0.003)[a]	0.365 (0.047)[a]

Data include 13 laboratories (329 slides) with two or more scorers; S.E.: standard error.
[a] Coefficient/S.E. ratio > 2.

Table 5
95% frequency intervals compatible with variance components estimated by the multilevel hierarchical model

	0 Gy	1 Gy	2 Gy
MNed	5.1–12.6	68.9–169.6	176.4–434.8
MN	5.3–14.4	76.8–208.6	221.6–602.1
NPB	1.4–12.4	5.8–51.0	12.9–113.0

Note: the results refer to frequency per 1000 BN cells.

and spot. Once these variables are considered, scorers from the same laboratory provide homogeneous readings. This pattern is similar for all the three endpoints.

The first part of Table 4 reports the log median ratio of fixed parameters representing the mean effect of dose and NDI. As would be expected, radiation dose is the strongest predictor of MNed frequency and despite its inclusion in the fixed part of the model this variable generates further heterogeneity, as shown by the random term. The information about variance components can be used to estimate the extent of the error due to the known sources of variability.

These results can be used to estimate ranges of MN frequency that are compatible with random variability due to major confounders. Using this approach we found that when the variability due to laboratory, staining method, and spot, are considered, the baseline frequency of MNed cells is between 5.1 and 12.6 MNed per 1000 BN cells. Similar estimates for all endpoints at different doses of irradiation are given in Table 5.

4. Discussion

The practical use of a biomarker of DNA damage in population monitoring is limited by (a) its reproducibility among laboratories when different laboratories attempt to compare their results obtained for their local populations, and (b) its reproducibility within the same laboratory when a population (which may include the same subjects within a population) is being studied repeatedly and longitudinally over an extended time-frame (e.g. years). The identification of factors affecting reproducibility in the CBMN assay, and a measure of the reproducibility, is essential for the determination of absolute MN frequency, accuracy in biological dosimetry of genotoxic exposure, and for the establishment of "normal" or "acceptable" base-line rates of chromosomal damage.

In spite of being provided slides prepared in identical manner from samples from the same cultures there was still a large extent of variation among laboratory scores for MN frequency. It is plausible that some of this variation could have been due to real differences between the cells in each sample. However, the fact that (a) each laboratory correctly ranked the slides according to radiation dose, and (b) those laboratories that scored low or high values for one dose also did so for the other doses, would tend to suggest that sampling differences are unlikely to be the main cause of the variability. The other possibility is that the interpretation of the scoring criteria provided, detailed as they were, ultimately depends on the subjective evaluation of each scorer and laboratory, so that some laboratories may have been more cautious than others with regard to accepting a binucleated cell with a micronucleus. The other practical limitation of identifying and visualising cells is the quality of the optics of the microscope used; we have not measured this difference between laboratories, although it is a potential source of variation that should be considered in future standardization studies.

We have observed considerable variation in the results of different scorers even within experienced laboratories. Assuming that scorers within a specific laboratory were using similar, or the same, microscopes it is still possible that there is variation in an individual's capacity to correctly identify a micronucleus or a nucleoplasmic bridge. Such discrimination would depend on the filters used, the precise focus, and the visual capacity to clearly identify micronuclear and nuclear boundaries. An important role is played by scorer's experience, as demonstrated by results in Table 4, which show that as the number of events to be scored increased with the increasing radiation dose, the ability to discriminate between exposed and control cells is more evident in those scorers with more experience. The observation of important inter-scorer variability fits well with the conclusion of Brown et al. [24] who also studied sources of variability in the CBMN assay using repeated scoring of slides from one culture and comparing scores of three different scorers. The analysis of their data showed a consistent performance of scorers on repeat counts and consistent, significant large differ-

ences in the results of different scorers. A measure of consistency of results from the same scorer is the coefficient of variation (CV). The results for base-line MN frequency CV for repeat counts by the same scorer values in this study (29%) are lower than those reported by Radack et al. [25] (52%) who, however, based their estimates on repeat measures from 100 BN cells instead of 1000 BN cells.

The extent of variation in the data for NPB was much greater than the variation observed for MN, as is clearly shown by the data in Fig. 3 and comparison of data in Figs. 5 and 6. This result was not unexpected given that only one laboratory was experienced in scoring this biomarker within the CBMN assay. Identification of NPB can be more difficult than identifying MN because the thickness of an NPB often can be much smaller than the diameter of an MN. In addition, NPB can be difficult to identify when nuclei in a BN cell are close to each other. Nevertheless, the great majority of the laboratories (31 out of 33) correctly identified the rank of the radiation dose.

The results in Table 4 clearly show that intra-laboratory variability is almost completely explained by the variables that we have included in the experiment—staining, dose and spot—and no inter-scorer variability remains once these are controlled. In comparison, as regards inter-laboratory variability, additional laboratory-linked features should be considered to fully understand this aspect. This high (and significant) degree of heterogeneity among laboratories is mainly based on the results of a limited number of laboratories, which scored greatly different frequencies of events from the majority. The use of random-effect models, which provides a better modeling of variance over the Poisson regression analysis, allowed the calculation of a range of 'reliable' values for different endpoints and levels of irradiation (Table 4). This estimate can be used as a reference for all laboratories when scoring the endpoints described here.

The results of this study raise important questions with respect to overcoming the inter-laboratory and inter-scorer variation, and indicate that even under close to optimal conditions the calibration of scorers and laboratories remains an important requirement. This leads immediately to the important goal of achieving a reliable procedure for calibrating scorers so that data from different scorers and laboratories can be corrected to values that are reliably close to the absolute/actual score. This goal could be achieved by using a common set of slides or, more practically, a set of digitized images made available on the HUMN project web site, together with detailed scoring instructions. This same training module could be used to train scorers to a standard method, thereby making more possible the achievement of comparable data between laboratories. However, the use of digitized images does not take into account inter-laboratory variation in the quality of microscope optics, indicating that ultimately the use of a standard slide set should also be used as a procedure for scorer calibration. Unraveling the effect of scoring criteria, the influence of microscope optics, and visual ability of the scorer, is a challenging scientific question in itself.

5. Conclusion

In conclusion, this inter-laboratory comparison has proved to be a fruitful exercise. It has verified that with respect to visual measurement of micronuclei we can confidently conclude that all laboratories correctly classified the slides according to radiation dose, and that dose of radiation accounted for most of the variation in the reported MNed cell frequency, while the importance of staining method was minimal. This study has estimated the expected CV for duplicate scores by the same scorer as being 29% for control cultures and 11–14% for cultures with MN frequencies that are an order of magnitude greater than that observed in control cultures. The reporting of intra-scorer CV values in the CBMN assay should become a standard requirement for all published studies, and large CV values should be carefully noted when evaluating the quality of the data. The other important product of this exercise was the development of a comprehensive set of scoring criteria for MN and NPB in human lymphocytes, as well as in other cell types that can be scored in the CBMN assay, together with a comprehensive set of line diagrams and photomicrographs for illustration, which are described in the accompanying paper [17]. Finally it is evident that the use of "calibration" or "control" slides is essential to reduce intra- and inter-laboratory variability and to allow comparison of MN frequency data across laboratories and populations.

Acknowledgements

The contribution of Ricardo Marcos, Alain Botta, Hubert Thierens, Kirsti Bredholt, Elena Szabova and others who have contributed to this study with their work is kindly acknowledged. We are indebted to Prof. Annibale Biggeri, Florence, Italy, for his assistance with the use of multilevel models. This work was supported by grants funded by the Associazione Italiana per la Ricerca sul Cancro (AIRC) and the European Union 5th FP (QLRT-2000-00628). Interested researchers are invited to collaborate and to get in touch with any member of the steering committee of the HUMN project (M.F., S.B., W.P.C., N.H., M.K.V., E.Z.). Further information can be found on the web site of the project: http://www.HUMN.org

References

[1] M. Fenech, N. Holland, W.P. Chang, E. Zeiger, S. Bonassi, The HUman MicroNucleus project—an international collaborative study on the use of the micronucleus technique for measuring DNA damage in humans, Mutat. Res. 428 (1999) 271–283.

[2] S. Bonassi, M. Fenech, C. Lando, Y.-P. Lin, M. Ceppi, W.P. Chang, N. Holland, M. Kirsch-Volders, E. Zeiger, S. Ban, R. Barale, M.P. Bigatti, C. Bolognesi, C. Jia, M. Di Giorgio, L.R. Ferguson, A. Fucic, O. Garcia Lima, P. Hrelia, A.P. Krishnaja, T.-K. Lee, L. Migliore, L. Mikhalevich, E. Mirkova, P. Mosesso, W.-U. Müller, Y. Odagiri, M.R. Scarfì, E. Szabova, I. Vorobtsova, A. Vral, A. Zijno, The HUman MicroNucleus project. International Data Base Comparison for results with the cytokinesis-block micronucleus assay in human lymphocytes. Part I. Effect of laboratory protocol, scoring criteria, and host factors on the frequency of micronuclei, Environ. Mol. Mutagen. 37 (2001) 31–45.

[3] M. Fenech, The in vitro micronucleus technique, Mutat. Res. 455 (2000) 81–95.

[4] M. Kirsch-Volders, T. Sofuni, M. Aardema, S. Albertini, D. Eastmond, M. Fenech, M. Ishidate, E. Lorge, H. Norppa, J. Surrales, W. von der Hude, A. Wakata, Report from the In Vitro Micronucleus Assay Working Group, Environ. Mol. Mutagen. 35 (2000) 167–172.

[5] K. Umegaki, M. Fenech, Cytokinesis-block micronucleus assay in WIL2-NS cells: a sensitive system to detect chromosomal damage induced by reactive oxygen species and activated human neutrophils, Mutagenesis 15 (2000) 261–269.

[6] M. Fenech, C. Aitken, J. Rinaldi, Folate, Vitamin B12, homocysteine status and DNA damage in young Australian adults, Carcinogenesis 19 (1998) 1163–1171.

[7] M. Kirsch-Volders, A. Elhajouji, E. Cundari, P. Van Hummelen, The in vitro micronucleus test: a multi-endpoint assay to detect simultaneously mitotic delay, apoptosis, chromosome breakage, chromosome loss and non-disjunction, Mutat. Res. 392 (1997) 19–30.

[8] W.P. Chang, B.F. Hwang, D. Wang, J.D. Wang, Cytogenetic effect of chronic low-dose, low-dose-rate gamma-radiation in residents of irradiated buildings, Lancet 350 (1997) 330–333.

[9] L.G. da Silva Augusto, S.R. Lieber, M.A. Ruiz, C.A. de Souza, Micronucleus monitoring to assess human occupational exposure to organochlorides, Environ. Mol. Mutagen. 29 (1997) 46–52.

[10] W.P. Chang, S.R. Lee, J. Tu, S.S. Hseu, Increased micronucleus formation in nurses with occupational nitrous oxide exposure in operating theaters, Environ. Mol. Mutagen. 27 (1996) 93–97.

[11] W.A. Anwar, S.I. Salama, M.M. el Serafy, S.A. Hemida, A.S. Hafez, Chromosomal aberrations and micronucleus frequency in nurses occupationally exposed to cytotoxic drugs, Mutagenesis 9 (1994) 315–317.

[12] A.D. Tates, F.J. van Dam, F.A. de Zwart, F. Darroudi, A.T. Natarajan, P. Rossner, K. Peterkova, K. Peltonen, N.A. Demopoulos, G. Stephanou, D. Vlachodimitropoulos, R.J. Sram, Biological effect monitoring in industrial workers from the Czech Republic exposed to low levels of butadiene, Toxicology 113 (1996) 91–99.

[13] H. Norppa, S. Luomahaara, H. Heikanen, S. Roth, M. Sorsa, L. Renzi, C. Lindholm, Micronucleus assay in lymphocytes as a tool to biomonitor human exposure to aneuploidogens and clastogens, Environ. Health Perspect. 101 (Suppl 3) (1993) 139–143.

[14] M. Fenech, A.A. Morley, Measurement of micronuclei in human lymphocytes, Mutat. Res. 148 (1985) 29–36.

[15] M. Fenech, A.A. Morley, Solutions to the kinetic problem in the micronucleus assay, Cytobios 43 (1985) 223–246.

[16] M. Fenech, A.A. Morley, Cytokinesis−block micronucleus method in human lymphocytes: effect of in vivo ageing and low−dose X−irradiation, Mutat. Res. 161 (1986) 193–198.

[17] M. Fenech, W.P. Chang, M. Kirsch-Volders, N. Holland, S. Bonassi, E. Zeiger, HUMN Project: detailed description of the scoring criteria for the cytokinesis-block micronucleus assay using isolated lymphocyte cultures, Mutat. Res. 534 (2003) 65–75.

[18] M. Fenech, The cytokinesis-block micronucleus technique: a detailed description of the method and its application to genotoxicity studies in human populations, Mutat. Res. 285 (1993) 35–44.

[19] M. Aitkin, D. Anderson, B. Francis, J. Hinde, Statistical Modelling in GLIM, Oxford Science Publications, Claredon Press, Oxford, 1990.

[20] S. Bonassi, V. Fontana, M. Ceppi, R. Barale, A. Biggeri, Analysis of correlated data in human biomonitoring studies. The case of high SCE frequency cells (HFC), Mutat. Res. 438 (1999) 13–21.

[21] H. Goldstein, Multilevel Statistical Models, Halsted, New York, 1995, p. 178.

[22] Generalized Linear Interactive Modeling (GLIM) System, Release 3.77, Numerical Algorithms Group, Oxford, UK, 1987.

[23] J. Rasbash, W. Browne, H. Golgstein, M. Yang, I. Plewis, M. Healy, G. Woodhouse, D. Draper, I. Langford, T. Lewis,

MLwiN Version 2.0. Multilevel Models Project Institute of Education, University of London, London, 1999.

[24] J.K. Brown, A. Williams, H.R. Withers, K.T. Ow, C. Gramacho, R. Grey, C. Amies, Sources of variability in the determination of micronuclei in irradiated peripheral blood lymphocytes, Mutat. Res. 389 (1997) 123–128.

[25] K.L. Radack, S.M. Pinney, G.K. Livingston, Sources of variability in the human lymphocyte micronucleus assay: a population-based study, Environ. Mol. Mutagen. 26 (1995) 26–36.