ELSEVIER

Mutation Research 463 (2000) 111–172

Reviews in Mutation Research

www.elsevier.com/locate/reviewsmr
Community address: www.elsevier.com/locate/mutres

IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans

Richard J. Albertini [a], Diana Anderson [b], George R. Douglas [c], Lars Hagmar [d], Kari Hemminki [e], Franco Merlo [f], A.T. Natarajan [g], Hannu Norppa [h], David E.G. Shuker [i], Raymond Tice [j], Michael D. Waters [k], Antero Aitio [l,*]

[a] *Genetic Toxicology Laboratory, The University of Vermont, Burlington, VT, USA*
[b] *TNO BIBRA International Ltd., Carshalton, Surrey, UK*
[c] *Environmental Health Directorate, Bureau of Chemical Hazards, Health and Welfare Canada, Ottawa, Ontario, Canada*
[d] *Department of Occupational and Environmental Medicine, Lund University, Lund, Sweden*
[e] *Department of Biosciences at Novum, Karolinska Institute, Huddinge, Sweden*
[f] *Environmental Epidemiology and Applied Biostatistics, National Cancer Institute, Genova, Italy*
[g] *Department of Radiation Genetics and Chemical Mutagenesis, Leiden University, Leiden, The Netherlands*
[h] *Laboratory of Molecular and Cellular Toxicology, Department of Industrial Hygiene and Toxicology, Finnish Institute of Occupational Health, Helsinki, Finland*
[i] *Biomonitoring and Molecular Interactions Section, MRC Toxicology Unit, University of Leicester, Leicester, UK*
[j] *Integrated Laboratory Systems, Inc., Research Triangle Park, NC, USA*
[k] *National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency, Research Triangle Park, USA*
[l] *Programme on Chemical Safety, World Health Organization, Geneva, Switzerland*

Abstract

The purpose of these guidelines is to provide concise guidance on the planning, performing and interpretation of studies to monitor groups or individuals exposed to genotoxic agents. Most human carcinogens are genotoxic but not all genotoxic agents have been shown to be carcinogenic in humans. Although the main interest in these studies is due to the association of genotoxicity with carcinogenicity, there is also an inherent interest in monitoring human genotoxicity independently of cancer as an endpoint.

The most often studied genotoxicity endpoints have been selected for inclusion in this document and they are structural and numerical chromosomal aberrations assessed using cytogenetic methods (classical chromosomal aberration analysis (CA), fluorescence in situ hybridisation (FISH), micronuclei (MN)); DNA damage (adducts, strand breaks, crosslinking, alkali-labile sites) assessed using bio-chemical/electrophoretic assays or sister chromatid exchanges (SCE); protein adducts; and hypoxanthine-guanine phosphoribosyltransferase (HPRT) mutations. The document does not consider germ cells or gene mutation assays other than HPRT or markers of oxidative stress, which have been applied on a more limited scale.

Keywords: Biomarkers; Biomonitoring; Genotoxic carcinogens; Human populations; Guideline; WHO

* Corresponding author.

1383-5742/00/$ - see front matter
PII: S1383-5742(00)00049-1

1. Preface

The monitoring of genotoxic effects of carcinogens in humans is increasingly applied for hazard identification or risk assessment purposes. It is recognised that this requires a multidisciplinary approach the lack of which has rendered many early studies difficult to interpret. Therefore, IPCS embarked on this project with the aim of producing concise guidelines on human monitoring studies, using a multidisciplinary approach.

The first drafts of the present documents were prepared by Drs. D. Anderson, and H. Norppa. On March 1–5, 1999, an expert group was convened in BIBRA International, to further work on the guidelines. This group consisted of Drs. R. Albertini, D. Anderson (Chairperson), G. Douglas, L. Hagmar, K. Hemminki, F. Merlo, A.T. Natarajan, H. Norppa (Rapporteur), D. Shuker, R. Tice, and M. Waters as participants, and Dr. A. Aitio as scientific secretary.

After an extensive discussion, it was agreed that the document should consist of a section of General Introduction (Section 2) followed by descriptions of assays for each individual endpoint as stand-alone documents. It was recognised that this will lead to repetition, but the advantage of stand-alone documents was considered to outweigh this disadvantage.

While it was recognised that diagrams on the mechanisms of formation for each endpoint might be helpful for some readers, it was felt that this is out of the scope of this document.

The document was sent for peer review to some 75 selected experts in the field. The following scientists sent their comments: Dr. J.W. Allen, Dr. Y. Barnett, Dr. C. Bolognesi, Dr. A. Brogger, Dr. J. Catalan, Dr. K.-S. Chia, Dr. R. Crebelli, Dr. F. Degrassi, Dr. P. Farmer, Dr. M. Fenech, Dr. L. Ferguson, Dr. S. Galloway, Dr. S. Garte, Dr. M.H.L Green, Dr. I.-L. Hansteen, Dr. M. Kirsch-Volders, Dr. D. Lovell, Dr. S. Madle, Dr. D. McGregor, Dr. L. Migliore, Dr. J. Mäki-Paakkanen, Dr. I. Nordenson, Dr. G. Obe, Dr. P. Olive, Dr. P. Ostrosky, Dr. K. Peltonen, Dr. N. Pearce, Dr. J. Preston, Dr. J. Ross, Dr. P. Sabbioni, Dr. M. Sorsa, Dr. G. Speit, Dr. R.J. Sram, Dr. E. Taioli, Dr. A.D. Tates, Dr. J. Tucker, Dr. H. Vainio, Dr. L. Vershaeve, Dr. J.B. Ward, Dr. J. Yaeger, Dr. E. Zeiger, Dr. A. Zijno.

All comments were carefully considered, and they led to an extensive rewriting of the guidelines by the members of the task group. The document was finally approved by the Chair and the Rapporteur in December, 1999. As is the standard with all IPCS documents, the views presented are those of the Task Group, and they do not necessarily represent the decisions of the stated policy of the United Nations Environment Programme, the International Labour Organization, or the World Health Organization.

2. General introduction

2.1. Purpose of the guidelines

The purpose of these guidelines is to provide concise guidance on the planning, performing and interpretation of studies to monitor groups or individuals exposed to genotoxic agents. Most human carcinogens are genotoxic but not all genotoxic agents have been shown to be carcinogenic in humans. Although the main interest in these studies is due to the association of genotoxicity with carcinogenicity, there is also an inherent interest in monitoring human genotoxicity independently of cancer as an endpoint.

The most often studied genotoxicity endpoints have been selected for inclusion in this document and they are structural and numerical chromosomal aberrations assessed using cytogenetic methods (classical chromosomal aberration analysis (CA), fluorescence in situ hybridisation (FISH), micronuclei (MN)); DNA damage (adducts, strand breaks, cross-linking, alkali-labile sites) assessed using biochemical/electrophoretic assays or sister chromatid exchanges (SCE); protein adducts; and hypoxanthine-guanine phosphoribosyltransferase (HPRT) mutations. The document does not consider germ cells or gene mutation assays other than HPRT or markers of oxidative stress, which have been applied on a more limited scale.

Using traditional criteria, the endpoints described herein can be divided into biomarkers of exposure, effect, and in some cases, susceptibility [1,2]. Distinction between these classes is not always defini-

tive, however, especially between biomarkers of exposure and effect. Interpretation of the results of biomarker studies, in the context of these guidelines, usually relates to the value of the endpoint, either as an indicator of cancer risk (i.e., biomarker of effect), or as a measure of exposure. For biomarkers of effect that can be proven to be predictors of disease per se, a reduction of biomarker values indicates a reduction in risk, but only if the biomarker is on a relevant pathogenic pathway. For biomarkers of exposure, reduction of a biomarker value due to intervention indicates a reduction in exposure, but not necessarily, a reduction of risk of all exposed. A biomarker of susceptibility is one which detects subjects who have a higher probability of an adverse effect due to individual variation.

This guideline also serves as a check-list for evaluating the methodology and results of completed studies. The main emphasis is on continuous exposures, such as those that occur in occupational/environmental settings, but an effort is also made to give guidance in situations where unexpected short-term exposures have occurred. These guidelines are not intended to be extensive reviews of each assay or how they are performed. For such information, the readers are referred to the original publications.

The guidelines aim at optimisation of the studies and cover both aspects that are strictly required and aspects that are desirable. Although an effort is made to distinguish between these two it is not always possible. Among those that are strictly necessary are ethical permission, defined exposed and referent populations, appropriate sample size, time of sampling, and referent population (subjects considered not to be exposed to the agent of interest, see Section 2.5), concurrent and identical handling of samples from both the exposed and referent subjects, and blinded analysis of coded samples.

2.2. General considerations relating to all genotoxicity endpoints

Inter-individual biological variability affects both biomarkers of exposure and effect. The mechanistic basis varies for different genotoxicity endpoints, and for some, it is not elucidated. Thus, although exposure to a variety of different carcinogens can be detected using these biomarkers, the causative agent cannot always be identified unequivocally, except sometimes in the case of specific DNA adducts. Furthermore, exposure biomarkers are usually more chemical-specific than effect biomarkers. In many cases, DNA and protein adducts and urine metabolites can be identified with a specific chemical. However, effect biomarkers, such as chromosomal aberrations or gene mutations are not chemical-specific, so the association with an exposure must be established by an independent measure. Exposure to carcinogens that act via nongenotoxic mechanisms is unlikely to be detected using any of the endpoints described in these guidelines. In such cases, monitoring should use other means (e.g., analysis of the chemical/metabolite in biological matrices). For some exposures and genotoxicity endpoints, a quantitative association between exposure and the biomarker has been demonstrated (e.g., ionising radiation, vinyl chloride, and benzene for chromosomal aberrations, ethylene oxide, aromatic amines, and acrylonitrile for haemoglobin adducts). Data are accumulating to support the concept that genotoxicity endpoints are predictors of human cancer risk (see Sections 4 and 7). Selection of the optimal assay and tissue for analysis for any particular exposure situation should be based, whenever possible, on exposure type and duration, exposed population characteristics, endpoint mechanism, and the expected target tissue [3]. Most of the time surrogate tissues rather than target tissues are studied for practical and methodological reasons. Where known, mechanistic information can be used in the selection of the appropriate assay to use for a specific exposure. The removal of damage occurs at different rates for different endpoints. Therefore, the selection of the endpoints also depends on the time at which the samples can be collected after exposure. Usually, the sooner the samples are taken, the better the chance of detecting the damage

2.3. Ethical considerations

Participation in a monitoring study should always be voluntary and in itself not pose a health risk. Participation or non-participation should never result in any form of discriminatory act against the subjects. A prerequisite for participation is a signed informed consent based on an adequate and under-

standable explanation of the intent of the study, any associated risks, and planned use for any samples, including possible future analyses of stored tissue. Personal information should be protected securely, and access to it should be limited. The participants should be informed in advance about the meaning of the possible results both on an individual and at a group level. Procedures should be established for the communication of possible ancillary findings of health relevance to all subjects. The results on a group level should be communicated in an understandable way to the participants and other relevant parties. The understanding and acceptance of the results generated may be facilitated by the presence of the study groups or their representatives in the planning process. Each participant has a right to be informed about his/her individual results, but such information should be accompanied by a competent interpretation. Individual results are to remain confidential. If the results of the monitoring program indicate health hazards, preventive action, i.e., prevention of exposure, at a group or even individual level needs to be addressed.

Ethical guidelines on human experimentation [4,5] and/or the applicable legislation should be abided. The monitoring study must be approved by the appropriate ethical authorities.

2.4. Sampling procedures

Blood and other human specimens should always be considered to be infectious material, and appropriate precautions for the prevention of contagion should be taken. The essential elements of safe laboratory practices include that (a) access to the laboratory is limited or restricted at the discretion of the laboratory director when experiments or work with cultures and specimens are in progress; (b) persons wash their hands after they handle viable materials and animals, after removing gloves, and before leaving the laboratory; (c) eating, drinking, smoking, handling contact lenses, and applying cosmetics are not permitted in the work area. Persons who wear contact lenses in laboratories should also wear goggles or a face shield. Food is stored outside the working area in cabinets or refrigerators designated for this purpose only. (d) Mouth pipetting is prohibited; mechanical pipetting devices are used; (e) all procedures are performed carefully to minimise the creation of aerosols; (f) work surfaces are decontaminated at least once a day and after any spill of viable material; (g) all cultures, stocks, and other regulated wastes are decontaminated before disposal by an approved decontamination method, such as autoclaving. Materials to be decontaminated outside of the immediate laboratory are to be placed in a durable, leakproof container and closed for transport form the laboratory. Materials to be decontaminated at off-site from the laboratory are packaged in accordance with applicable local, state, and federal regulations, before removal from the facility; (h) an insect and rodent control program is in effect [6]. More detailed guidelines on biohazard prevention are available [7].

Biological samples should be collected according to the protocol of the assay employed taking into account the biological response time of the end point under investigation. If the biological response time of the parameter measured is not known, then repeated sampling may improve the likelihood of detecting a positive response. If studies to be performed include several end points, separate samples, different sample preparation procedures and storage conditions may be required (e.g., if plasma proteins are to be measured, proteolysis must be inhibited; if genotyping is to be performed, heparin must be avoided). A transportation plan for samples must be developed and field tested to ensure its workability. Wherever possible, storage of specimens in a repository is recommended. Considerations should be given to the stability of the stored sample for the endpoint to be studied. A prerequisite for future studies using stored specimens is a signed informed consent from each donor. The form should indicate what future studies are contemplated.

2.5. Size and characteristics of studied populations

A study protocol must be specified in advance taking into account the composition of the referent and exposed groups. It should clearly state (a) primary and secondary objectives, (b) the size of the effect(s) to be considered biologically important (c) ideally, the statistical methods for data analysis. When embarking on a study, the desired power should be defined a priori. The statistical power of the study is characterised by its ability to detect a

true effect of a given magnitude in a statistically significant way, as defined in the study protocol. (e.g., type I or alpha error = 0.05). The three important determinants of the power are (1) the magnitude of the expected (assumed) effect of exposure to a genotoxic agent, (2) the variability of the biomarker (i.e., the variability of the measured endpoint), and (3) the size of the study populations. The magnitude of the expected effect (e.g., absolute difference or relative effect) is a key aspect in determining the biological importance of the monitoring study. The variability of the biomarker may be due to both analytical factors and inter- and intra-individual variation and sampling errors. If the variability of the biomarker is not well known, a pilot study might be needed in order to optimise the etiological study. An effective way of increasing the power is to decrease the laboratory variation (i.e., increase assay precision) by controlling for measurement errors The adequate number of exposed and non-exposed subjects to be monitored (i.e., sample size) should be calculated, keeping in mind the smallest effect that is of biological relevance — which has to be considered for each end point. An 80% power is often considered acceptable, even though this allows a 20% chance of missing a true positive response (type II or beta error).

There is user-friendly software that can be used for sample size or power calculations [8–10]. A practical example will illustrate power calculations: assume that in the referent group, the mean value for the biomarker equals 2.2 and the standard deviation is 2.1. If a statistical power of 80% is required for significantly detecting (i.e., obtaining a two-tailed p-value of < 0.05), a difference between the exposed and referent groups that corresponds to a 50% higher mean value among the exposed ($3.3 - 2.2 = 1.1$), 58 subjects are needed in each group. Alternatively, if the relevant difference corresponds to a 30% higher mean value among the exposed ($2.86 - 2.2 = 0.66$), 160 subjects are needed in each group. When considering the sample size of a study it is important to consider that the numbers have to be large enough to allow for failure of compliance, unacceptable controls or loss of samples due to a failure of transport or laboratory handling. In addition, the power calculations represent an ideal non-biased study design. In designing a study, those factors that are known to influence the level of each biomarker should be controlled for. This means that referent subjects have to be selected while accounting for relevant matching factors, such as age, gender, ethnicity, socio-economic status, disease status, and smoking habits. Individual matching or frequency matching can be applied [11]. Expert statistical advice should be taken to take these factors into account.

The number of referents for each exposed subject should, whenever possible, be optimised with respect to the statistical power of the study. However, it is appreciated that this is not always possible due to, for example, cost constraints.

The term ''referent'' will be used throughout this document to refer to subjects considered not to be exposed to the agent of interest. This term is used rather than the term ''control'' since the latter generally refers in human studies to the absence of a disease state.

The genotype of some xenobiotic metabolising enzymes have been demonstrated to affect either the baseline level or chemically induced level of DNA adducts, SCEs, and chromosomal aberrations in humans. However, study subjects can only infrequently be pre-selected on the basis of their genotype/phenotype for metabolising enzymes. Thus, it may be useful to control for genetic polymorphisms by employing multivariate statistical models. It should be noted, however, that the expected prevalences of these factors will affect the study power and should therefore be taken into account in the study design.

Sampling should be performed during a relevant time period with respect to exposure and the genotoxic endpoint of interest. Most of the genotoxic endpoints discussed here involve an assessment of DNA damage either directly (e.g., with biochemical techniques) or indirectly (e.g., SCE, chromosomal aberrations, MN). Much DNA damage is short-lived (less than two days) and samples should be collected as quickly as possible following acute exposures or the termination of a chronic exposure. Protein adducts are more persistent (up to weeks or even a few months depending on the turnover rate of the protein), and repeated samplings on individuals can be used for dosimetry measurements. Mutagenic/genotoxic end points on the other hand, have a rather long biological half-live (related more to cell lifespan)

and, therefore, timing of specimen collection is seldom a question of hours or a few days. In accidental acute exposure situations, there is a general need to quickly collect samples, with the optimal sample time depending on the genotoxic endpoint of interest. Furthermore, in such situations, the number of exposed subjects available for study may be limited and the exposure assessments less thorough than that possible for chronic exposures. Therefore, the means to reach a desired study power may be limited, with respect to study population size. In such situations, it can be worthwhile to invest even more resources than usual in decreasing technical variability (i.e., increase the number of replicate analyses). However, even if a decrease in technical variability is possible, the study design will still have a low power, and careful considerations must be made with respect to whether the study should be performed or not.

Samples should always be collected from exposed subjects and referents concurrently and not successively, and the laboratory analyses performed in a random order [12]. This will minimise misclassification due to drift in assay performance. The samples must be coded and assayed without knowledge of subject identity with respect to exposure status to guarantee identical handling of samples from both exposed and referent subjects.

2.6. Assessment of individual exposure

Assessments — qualitative and quantitative — should be done for each monitored subject (both ''exposed'' and referents) with respect to the exposure of interest and other relevant chemical or physical agents. Intensity, duration, and pattern (e.g., peak exposures) should be considered. For long-term exposures, the stability of the exposure level over time should be determined. Efforts should be made to measure peak exposures. They may be important, e.g., due to overloading of defence mechanisms, in producing genotoxic effects and cancer [13]. The time period of relevance for the exposure assessment and the length of time between exposure and specimen collection is dependent on the outcome measure of interest. With respect to occupational exposures, the primary source of information is historical industrial hygiene data and working history. It is useful to ask participants to keep a diary of their activities and unusual events that may alter exposure during the study period. Levels of the chemicals or their metabolites in blood or urine may be useful in the exposure assessment but these measurements usually reflect only a very recent period of exposure.

2.7. Assessment of potential confounders and effect modifiers

In addition to the exposure of interest and other relevant chemical and physical exposures, other factors may affect the outcome measures of interest. The importance of such potential confounders or effect modifiers may vary for different endpoints. Data on potential confounders or effect modifiers need to be collected by, for example, questionnaires, interviews, or analyses (e.g., ethnic group, age, sex, tobacco smoking, drug usage, exposure to X-rays, chemical exposure at home, genetic polymorphisms).

2.8. Statistical analysis

Comparisons should be conducted between the exposed group and the referent group and, whenever possible, considering exposure–response associations or time trends according to the statistical plan defined in the study protocol. Statistical analysis may be carried out using (1) a hypothesis testing approach (Is there an effect?) or (2) an estimation approach (How big is the effect?).

A positive result, using hypothesis testing, depends on the magnitude of the observed effect, on a lack of systematic errors in study design and the degree of statistical significance obtained [14]. A prerequisite for concluding that the result of a study is negative is not only a sufficient study power but also a lack of systematic errors. The alternative approach, based on estimation approach, is to provide an estimate of the size of an effect together with appropriate confidence intervals [15]. These two approaches can complement each other.

For individual cell-based assays, when considering the number of cells to be analysed for each subject, the basal frequency of the endpoint should be considered. Wherever possible, the proportion of zero values should be kept as low as technically feasible. When applicable, zero values can be trans-

formed by different ways, i.e., by applying the average square root transformation $[(\sqrt{x} + \sqrt{x+1})/2]$ [16]. Another possibility is to assign the zero values the value of a fraction of the detection limit.

The applicability of statistical methods is often limited by the number of observations (see Section 2.5). Depending on the shape of distribution of the measured end points, medians and percentile distributions or means and standard deviations should be used to describe the results. Graphical representation of data (e.g., box and whisker plots, scatter plots) may be useful tools. The statistical analysis plan should be honoured and in hypothesis testing, the assumptions should be verified. If the assumptions underlying the statistical methods are violated, alternative statistical approach may need to be used. There is a wide range of statistical methods available for the analysis of genotoxicity endpoints. One consideration in choosing the method is the distribution of the endpoint among individuals. If the distribution is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, if appropriate data transformation is feasible, parametric methods based upon normal distribution can also be used.

In the estimation approach, the relative and absolute effects of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion derived from the hypothesis testing approach by estimating the precision of the effects seen in the study.

Stratified analyses can sometimes be used for exploring the effect modification or confounding by factors such as age, sex, smoking, other exposures, etc. However, only seldom will the sample size of the study allow a full evaluation of confounding and effect modification by stratified analyses. Multivariate statistical methods are valuable tools in such situations.

2.9. Publishing and archiving of data

A comprehensive report, as described separately for each endpoint, should be prepared. The studies should be published in the open literature. Parts of the reports are generally not accepted for publication by the journals. Raw data should be archived according to national and local requirements. Where the questionnaire cannot be published, it should be archived and made available upon request to bona fide scientists.

2.10. Quality of data

All procedural protocols should be adopted or developed in advance. They should be used with realistic trial samples before being applied to actual study materials. Considerations should be given to the use of existing external quality control schemes and to establishing such procedures if not available. The internal quality assurance procedures adopted and the uncertainties of the methods must be described. Some laboratories use the expected difference between smokers and nonsmokers or young and old subjects as an indicator of analytical quality or as a positive control. Some equally reputable laboratories, however, do not see such differences and the usefulness of these factors as means of quality assurance is not universal. For all genotoxicity endpoints, the unprocessed specimens are prone to deterioration if not properly stored. Traditional quality assurance measures are not likely to identify changes in sample characteristics taking place before the analysis. Therefore, sample tracking and verification of sample integrity are of prime importance. In the assays based on the collection of data using microscopic analysis, and where multiple scorers are used, an analysis of interscorer variability should be conducted. This mandates that data be collected in such a manner that scorer identity is maintained. As a routine, standard operating procedures should be developed and followed, and Principles of Good Laboratory Practices [17] (e.g., Good Epidemiological Practices [18], and Good Clinical Practices [19] should be adhered to, whenever feasible.

2.11. Literature

1. National Research Council (USA), Biologic markers in reproductive toxicology, National Academy Press, Washington, DC, 1989, 395 pp.
2. International Programme on Chemical Safety (IPCS), Biomarkers and risk assessment: con-

cepts and principles, Environmental Health Criteria, 155, World Health Organization, Geneva, Switzerland, 1993, 82 pp.
3. A.D. Tates, T. Grummt, F.J. van Dam, F. deZwart, F.J. Kasper, R. Rothe, et al., Measurement of frequencies of HPRT mutants, chromosomal aberrations, micronuclei, sister-chromatid exchanges and cells with high frequencies of SCEs in styrene/dichloromethane-exposed workers, Mutat. Res. 313 (1994) 249–262.
4. Council for International Organizations of Medical Sciences, International Ethical Guidelines for Biomedical Research Involving Human Subjects, Geneva, 1993.
5. World Medical Association, Declaration of Helsinki, JAMA, 277 (1997) 925–926.
6. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health. Biosafety in Microbiological and Biomedical Laboratories. HHS Publication No (CDC) 93-8395, 3rd edn. (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H. Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices, 2nd edn., ASM Press, Washington DC, 1995.
7. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
8. Dixon Associates, nQuery Advisor® Version 3.0, Los Angeles, CA.
9. Cytel Software, Egret-Siz. Cambridge, MA.
10. SPSS, Sample Power 1.2, Chicago, IL.
11. K.J. Rothman, Matching, in: K.J. Rothman, S. Greenland (Eds.), Modern Epidemiology, 2nd edn., Lippincott-Raven Publishers, Philadelphia, PA, 1998, pp. 147–162.
12. P. Vineis, P.A. Schulte, J. Vogt, Technical variability in laboratory data, in: P.A. Schulte, F.P. Perera (Eds.), Molecular Epidemiology: Principles and Practices, Academic Press, San Diego, 1993, pp. 109–135.
13. E.J. Calabrese, R.B. Blain, The Single Exposure Carcinogen Database: assessing the circumstances under which a single exposure to a carcinogen can cause cancer. Toxicol. Sci. 50 (1999) 169–185.
14. K.J. Rothman, S. Greenland, Precision and Validity in Epidemiologic Studies, in: K.J. Rothman, S. Greenland (Eds.), Modern Epidemiology, 2nd edn., Lippincott-Raven Publishers, Philadelphia, PA, 1998, pp. 115–134.
15. M.J. Gardner, D.G. Altman, Confidence intervals rather than P values: estimation rather than hypothesis testing, Br. Med. J. 292 (1988) 746–750.
16. E.B. Whorton, Some experimental design and analysis considerations for cytogenetic studies, Environ. Mutagen. 7 (Suppl. 4) (1985) 9–15.
17. Organisation for Economic Co-operation and Development, Council Decision — Recommendation on compliance with principles of good laboratory practice, Adopted by the Council at its 717th session on 2nd October 1989.
18. Ethics in Epidemiology, J. Clin. Epidemiol. 44 (Suppl. 44) (1991) 1S–151S.
19. J. Idänpään-Heikkilä, WHO guidelines for good clinical practice (GCP) for pharmaceutical products: the responsibilities of the investigator, Ann. Med. 26 (1994) 89–94.

3. Protein adducts

3.1. Introduction

3.1.1. Purpose of the guideline

To provide concise guidance for various methodologies for the detection and quantitation of protein adducts in humans. This guideline must be read in conjunction with the General Introduction (Section 2), which contains information on such issues as study design, ethics, and appropriate statistical considerations.

3.1.2. Principles of the assay

Many different electrophilic chemicals or their metabolites are known to react with amino acid side chains in haemoglobin and albumin [1]. The principal sites of reaction are at cysteine, histidine and, for normal (i.e., alpha, beta) haemoglobin, N-terminal valine. As human erythrocytes have a lifetime of about 120 days, measurement of haemoglobin adducts can provide information on exposures occur-

ring several weeks to months prior to the blood sample being taken [2]. The half-life of albumin in humans is somewhat shorter at about 28 days [3].

3.1.3. Definitions

A protein adduct is a chemical entity bound covalently to protein.

3.1.4. Significance of the endpoint and application in risk assessment

Most human carcinogens form DNA adducts and, for some of them, protein adducts have also been characterised [1,4]. Protein adducts are used primarily as biomarkers of exposure to genotoxic agents, i.e., identification and quantification of exposure. For a few agents, such as ethylene oxide, the relationship between protein adducts and DNA adducts has been defined and the former has been used in risk assessment [1,5].

3.2. Population characteristics

Important characteristics to consider for the exposed and referent populations are discussed in Section 2. As presented in this section, specific information on lifestyle (e.g., smoking, recreational/medical drug usage), health status, and endogenous factors, such as age and sex, for each subject may be useful. When feasible, genetic polymorphisms associated with the metabolism of the chemical under study should be identified. A concurrent referent population is always required.

As discussed in Section 2, the number of subjects needed for a particular study design depends, for a desired power, on the magnitude of the expected effect and the known variability of the measured endpoint. Appropriate software programs are available to assist the investigator in making these calculations (see Section 2). Published information on population size calculations for protein adducts is not available.

Where the number of subjects to evaluate cannot be determined prior to study implementation or in situations where the appropriate number of subjects are not available, the uncertainties, qualitatively and quantitatively, of the result should be considered in the interpretation.

3.3. Sample timing

The optimal collection time of blood for an assessment of protein adducts is during a long-term chronic exposure when the induction and loss of protein adducts is presumed to be at steady-state. There appears to be no active repair of protein adducts and they disappear at approximately the same rate as turnover of the respective proteins. However, there may be exceptions and the chemical stability of the adduct needs to be verified. Thus, in contrast to DNA adducts, protein adducts can be used to detect exposures that occurred days, weeks, and, in some cases, months prior to sample collection.

3.4. Methods

3.4.1. Quality assurance

A system of recording the transport and storage of samples should be established to ensure that the integrity of each sample is maintained. Samples should not be stored in close proximity to reactive chemicals and preferably should be stored in dedicated refrigerators/freezers. Reactive chemicals that are volatile may lead to cross-contamination; to avoid such contamination, air-tight individual packaging of each sample may be necessary. The use of coded samples where analysis is conducted without knowledge of exposure information is critical for eliminating potential bias. Protein containing known amounts of adducts should be used as quality assurance samples each time a series of samples is analysed, and standardised protocols should be followed.

3.4.2. Blood collection

Because of the risk for infectious diseases (hepatitis, HIV), appropriate precautions must always be followed when handling blood samples ([6,7]; see Section 2).

3.4.3. Description of methods

3.4.3.1. Haemoglobin adducts. Red blood cells are collected from whole blood by centrifugation and lysed to release haemoglobin [8]. Erythrocytes can be stored for future analysis of haemoglobin adducts when blood is sampled for other purposes [9]. Globin is precipitated by treatment with hydrochloric acid/acetone, washed with acetone, and finally washed with diethyl ether. Dry globin can be stored at 4°C for several months prior to analysis. The adducts are analysed by derivatisation and gas chromatography-mass spectrometry (GC-MS) for *S*-alkylcysteines and *N*-alkyl histidines, or a modified Edman reaction followed by GC-MS for *N*-alkylvalines. As most agents that modify cysteine and histidine also react with the N-terminal valine, the modified Edman procedure has become widely used [5,10]. As an alternative method for aromatic compounds, mild hydrolysis can be used to release the bound compound, which is then analysed by GC-MS [4].

Modified Edman method for analysis of N-terminal valines. Globin samples in formamide are treated with pentafluorophenylisothiocyanate. The reaction mixture is quenched with bicarbonate solution and the pentafluorophenylthiohydantoin derivatives of *N*-alkylvalines are extracted into toluene and analysed by GC-MS. Individual *N*-alkylvalines are quantitated by comparison with stable isotope labelled internal standards [11].

3.4.3.2. Albumin adducts. Albumin is extracted from serum by affinity chromatography or ammonium sulphate precipitation and adducts are detected by one of two methods: (1) immunochemical detection of the intact protein adduct (e.g., aflatoxin B_1) [12], (2) hydrolysis of the adducted chemical and analysis of the released species (e.g., 4-aminobiphenyl or *S*-phenylcysteine) [3].

3.4.4. Sources of assay variation

A lack of uniform collection, shipping, processing, and storage of samples will increase experimental variability.

3.4.5. Statistical analysis

As described in Section 2, data should be evaluated according to the statistical plan of the study. For each subject, results should be expressed as moles adduct per gram normal protein showing mean values ($\pm$standard deviation) or medians ($\pm$percentile distribution) of duplicate to triplicate analyses of samples. The choice of statistical method to use may depend on the distribution of the data. If the distribution of the data is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, parametric methods based upon normal distribution can also be used, when appropriate, after relevant data transformation. When needed, zero values can be transformed by different ways [13] or by assigning them the value of a fraction of the detection limit. Statistical advice may be needed to determine whether parametric or non-parametric methods are more appropriate for a particular study. The relative and absolute effects of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion by estimating the precision of the effects seen in the study. Stratified analyses can be used for exploring the effect modification by factors such as age, sex, smoking, other exposures, etc. More sophisticated multivariate methods may be necessary for investigating the effect of factors that either may modify or be confounded with the effects. Wherever possible, the analysis should be based on individual exposure information. Consideration for the effect of multiple comparisons on the false positive rate may be included in the statistical analysis.

3.4.6. Interpretation of results

Criteria for identifying a positive response (i.e., increased levels of adducts in an exposed group compared to the referent population; a dose response relationship if variable exposure levels have been established) must be identified prior to implementing the study.

Generally, a statistically significantly higher level of protein adducts in the exposed population in comparison to the unexposed referents is considered a positive result. A positive result is supported by a dose–response relationship when individuals with different exposure levels are considered. Maximum reliability that an exposure result in increased levels

of protein adducts requires reproducible results in independent studies. It is appreciated that independent studies are not always possible (e.g., in the case of accidents).

The lack of a statistically significant increase in protein adducts in a particular study indicates, only, that under the exposure conditions evaluated and for the calculated power of the study, the exposure did not result in a increase in adducts in the protein evaluated.

The results of the quality assurance analyses are used to demonstrate the adequacy of the methodology for detecting protein adducts.

3.5. Report

All data should be presented in tabular and/or graphical form, and include all observed results (see below). This will allow comparisons between laboratories, re-analyses of results by outside reviewers, and provide a database for baseline background values and for exposures to physical and chemical agents. All results of the study should be reported with none omitted, and the definition and handling of ''outliers'' described.

The study report should include the following minimal information:

- copy of the questionnaire used to determine personal characteristics
- statement of ethical approval
- characteristics of the exposed and referent populations
- exposure information and analytical methods used
- blood collection procedures, including timing in relationship to exposure, transportation, storage, and chain-of-custody
- method for coding samples prior to analysis
- methods for protein extraction and quantitation
- method used to generate standard curves
- representative printouts of raw data and standard curves
- chromatographic and mass spectral conditions
- representative GC-MS chromatograms
- tabulated results for each subject
- statistical methods
- discussion and interpretation of results
- internal and external quality assurance procedures

3.6. Literature cited

1. M. Törnqvist, L. Ehrenberg, On cancer risk estimation of urban air pollution, Environ. Health Perspect. 102 (Suppl. 4) (1994) 173–181.
2. M. Törnqvist, J. Mowrer, S. Jensen, L. Ehrenberg, Monitoring of environmental cancer initiators through haemoglobin adducts by a modified Edman degradation method., Anal. Biochem. 154 (1986) 255–266.
3. S.R. Tannenbaum, P.L. Skipper, J.S. Wishnok, W.G. Stillwell, B.W. Day, K. Taghizadeh, Characterisation of various classes of protein adducts, Environ. Health Perspect. 99 (1993) 51–56.
4. P. Farmer, Monitoring of human exposure to carcinogens through DNA and protein adduct determination, Toxicol. Lett. 82–83 (1995).
5. M. Törnqvist, H.H. Landin, Haemoglobin adducts for in vivo dose monitoring and cancer risk estimation, J. Occup. Environ. Med. 37 (1995) 1077–1085.
6. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
7. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health, Biosafety in Microbiological and Biomedical Laboratories. HHS Publication No (CDC) 93–8395, 3rd edn (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H. Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices, 2nd edn., ASM Press, Washington DC, 1995.
8. M. Sorsa, K. Peltonen, D. Anderson, N.A. Demopoulos, H.-G. Neumann, S. Osterman-Golkar, Assessment of environmental and occupational exposures to butadiene as a model for risk estimation of petrochemical emissions, Mutagenesis 11 (1996) 9–17.
9. A.D. Tates, T. Grummt, M. Törnqvist, P. Farmer, F.J. Van Dam, H. van Mossel et al., Biological

and chemical monitoring of occupational exposure to ethylene oxide, Mutat. Res. 250 (1991) 483–497.
10. S. Osterman-Golkar, J.A. Bond, Biomonitoring of 1,3-butadiene and related compounds, Environ. Health Perspect. 104 (Suppl. 5) (1996) 907–915.
11. M. Törnqvist, A.-L. Magnusson, P.B. Farmer, Y.-S. Tang, A.M. Jeffrey, L. Wazneh et al., Ring test for low levels of *N*-(2-hydrosyethyl)valine in human hemoglobin, Anal. Biochem. 203 (1992) 357–360.
12. J.D. Groopman, T.W. Kensler, The light at the end of the tunnel for chemical-specific biomarkers: daylight or headlight? Carcinogenesis 20 (1999) 1–11.
13. E.B. Whorton, Some experimental design and analysis considerations for cytogenetic studies, Environ. Mutagen. 7 (Suppl. 4) (1985) 9–15.

4. DNA adducts

4.1. Introduction

4.1.1. Purpose of the guideline

To provide concise guidance for various methodologies for the detection and quantitation of DNA adducts in human tissues. This guideline must be read in conjunction with the General Introduction (Section 2), which contains information on such issues as study design, ethics, and appropriate statistical considerations.

4.1.2. Principles of the assays

DNA is isolated from peripheral lymphocytes, whole blood, or tissues sampled from exposed and referent individuals. DNA adducts are detected at the level of modified bases, deoxynucleosides, or oligonucleotides by a range of physico-chemical or immunochemical techniques (see below). In some cases, DNA repair results in the excretion of a modified base in urine and this can be used as a measure of DNA damage [1].

There are two types of techniques that are currently used for the analysis of DNA damage. *Chemical-specific techniques* rely on knowledge of the structure and properties of the particular DNA adducts to be detected and are often oriented towards the measurement of DNA damage in situations where the exposure circumstance is well characterised. *Non-specific techniques* demand only a minimum requirement, such as the fact that DNA is covalently modified in such a way that it perturbs the function of an enzyme as in the case of the nuclease P1 version of ^{32}P-postlabelling. The chemical identities of the DNA adducts formed are typically not determined.

4.1.3. Definitions

A DNA adduct is a chemical entity bound covalently to DNA [2].

4.1.4. Significance of the endpoint and application in risk assessment

DNA adducts are used primarily as biomarkers of exposure to genotoxic agents. Most human carcinogens are genotoxic and DNA adducts have been characterised for many of them [2]. The relationship between adducts, their persistence and repair, and mutagenic endpoints are complex. Most DNA adduct measurements are global measures of specific DNA alterations or changes in genomic DNA and often the DNA is extracted from heterogeneous cell populations. However, there is epidemiological evidence that some DNA adducts can be predictive of cancer risk (e.g., aflatoxin B1 and hepatocellular cancer [3] and can be used as short term indicators of the effects of intervention [4]. In contrast, there is also experimental evidence that levels of DNA adducts in a particular tissue correlate in general with exposure but the relative quantity of adducts per tissue/organ type is not necessarily predictive of subsequent tumour formation in that tissue [5,6].

4.2. Population characteristics

Important characteristics to consider for the exposed and referent populations are discussed in Section 2. As presented in this section, specific information on lifestyle (e.g., smoking, recreational/medical drug usage), health status, and endogenous factors, such as age and sex, need to be obtained for each

subject. Wherever possible, individuals with genetic polymorphisms associated with increased susceptibility to genotoxic agents should be identified. Data demonstrating that these factors modulate the frequency of adducts in human tissues are lacking although there are examples of associations between DNA adduct levels and metabolic phenotype [7–9]. A concurrent referent population is always required.

As discussed in the General Introduction, the number of subjects needed for a particular study design depends, for a desired power, on the magnitude of the expected effect and the known variability of the measured endpoint. Appropriate software programs are available to assist the investigator in making these calculations (see Section 2). However, there is no published information on population size calculations for studies involving DNA adducts.

Where the number of subjects to evaluate cannot be determined prior to study implementation or in situations where the appropriate number of subjects are not available, the precision in the risk estimate should be considered in evaluating the results.

4.3. Sample timing

The optimal collection time for any human tissue is during a long-term exposure when the induction and repair of DNA damage is presumed to be at steady-state. Also, for some exposures, there is evidence that measurement in urine of adducts excised from DNA during excision repair processes can be indicative of cancer risk [3,10]. For the sampling of tissues after an acute exposure or after termination of a chronic exposure, the optimal collection time for detecting DNA adducts is usually within a few hours to days of exposure termination. The level of DNA damage decreases with increasing time between exposure and sampling due to DNA repair processes, the loss of heavily damaged cells through apoptosis or necrosis, and to cell turnover. Some types of DNA adducts (e.g., *N*-7-guanine and *N*-3-adenine derivatives) are chemically unstable and decompose with a half-life of about 3–5 days [11]. Most types of DNA adduct are actively repaired by enzymatic processes [12]. The half-life of UV-induced cyclobutane dimers is approximately 15 h in human skin [13], while the half-lives of some aromatic DNA adducts in peripheral lymphocytes are assumed to be > 1 month [14]. Thus, in general, the levels of adducts are usually indicative of recent exposures.

4.4. Methods

4.4.1. Quality assurance

A system of recording the transport and storage of samples should be established to ensure that the integrity of each sample is maintained. Samples should not be stored in close proximity to reactive chemicals and preferably should be stored in dedicated refrigerators/freezers. The use of coded samples where analysis is conducted without knowledge of exposure information is critical for eliminating potential bias. DNA containing known amounts of adducts should be used as quality assurance samples each time a series of samples is analysed and standardised protocols should be followed (e.g., for ^{32}P-postlabelling see Ref. [15]). Analyses are performed in duplicate or triplicate.

4.4.2. Sample collection

Because of the risk for infectious diseases (hepatitis, HIV), appropriate precautions must always be followed when handling human tissue samples [16,17].

4.4.3. Description of methods

4.4.3.1. DNA purification. DNA is extracted from peripheral lymphocytes, whole blood, or tissues by either phenol/chloroform extraction, salting out with or without polyamine precipitation [18,19] or using proprietary solid phase extraction kits. The purity of isolated DNA should be assessed by an appropriate technique, such as UV spectroscopy, gel electrophoresis or HPLC. DNA, which is not adequately pure (i.e., containing RNA or protein), should be repurified.

4.4.3.2. ^{32}P-postlabelling. Many variations of the postlabelling assay have been developed to detect certain classes of adducts. The main points of variation include (1) selection of enzymes for digestion of

DNA, (2) enrichment methods for adducts, (3) labelling conditions, and (4) separation of adducts [13,15,20,21]. In the original method, samples of the purified DNA are hydrolysed to the constituent 3′-mononucleotides using micrococcal endonuclease and spleen exonuclease. The adducts have to be enriched by some method (see below) because normal nucleotides are in vast molar excess of adducts (1 μg DNA = 16 nmol; radioactive ATP is used in pmol quantities). Adducted nucleotides in the DNA digest are labelled at the 5′ position with ^{32}P-labelled γ-ATP and T4 polynucleotide kinase and the labelled nucleoside 3′,5′-diphosphates separated by PEI-cellulose thin-layer chromatography or high performance liquid chromatography (HPLC). If small DNA adducts are analysed, dephosphorylation at the 3′ end using P1 enzyme substantially enhances the resolution of structurally different adducts [1]. Quantification is carried out by analysing synthetic standard compounds in parallel experiments. Depending on the assay and its associated costs, the whole procedure should be repeated once or twice.

The resolved adducts are located on the PEI-cellulose chromatograms by autoradiography using X-ray film and intensifying screens or visualised with storage phosphor or other direct two-dimensional radioactivity detectors, or by analysis of phosphor image or detector output [18,19]. The individual adduct spots are quantified by scintillation counting. In the HPLC methods, flow-through radioactivity detection can be used and adduct fractions quantified by peak area [13,20]. Prior to postlabelling, enrichment of small adducts has been achieved using HPLC and of bulky adducts using *n*-butanol extraction [15,18,19]. Increased adduct labelling can also be obtained by the selective enzymatic dephosphorylation of normal (unadducted) nucleotides by nuclease P1 treatment prior to post-labelling. This latter method may result in an underestimation of adduct levels if the adducted nucleotides are only partially resistant to nuclease P1 dephosphorylation.

A significant advantage of ^{32}P-postlabelling over many other methods for the determination of DNA damage is that the method is capable of detecting and quantitating modifications of unknown structure, as exemplified by the I (indigenous) compounds [22]. The ^{32}P-postlabelling technique is also the most widely used method to detect DNA adducts.

4.4.3.3. Mass spectrometry (MS). DNA is broken down to bases, nucleosides, nucleotides, or oligonucleotides by chemical or enzymatic digestion and adducts are separated by chromatography prior to detection and quantitation by appropriate MS techniques. The presence of adducts is detected by increases in the molecular weight over the normal bases; the nature of the adduct may be elucidated by further mass spectral analyses. Quantitation is usually by comparison to stable isotope labelled internal standards [23,24].

4.4.3.4. Radioimmunoassay. A constant amount of radioactively labelled DNA adduct (tracer) is mixed in the assay tubes with a constant amount of antibody sufficient to bind 50% of the tracer. Unknown samples (i.e., containing unlabelled DNA adduct) displace a fraction of the bound tracer. Upon precipitation of the antibody bound tracer by a variety of procedures, the fraction of radioactive tracer remaining is determined by scintillation counting [25–28].

4.4.3.5. ELISA. A variety of enzyme immunoassay procedures have been described and there is no general procedure applicable to all DNA adducts. The salient features of an ELISA procedure for DNA adducts are as follows: a known amount of adduct-containing DNA or DNA-adduct bound to protein is adsorbed onto the surface of microtitre wells or polystyrene tubes. The samples containing unknown amounts of DNA adduct are added to the wells/tubes followed by a predetermined constant concentration of antibody. Upon washing the wells/tubes with buffer, unbound antibody and adduct–antibody complexes are washed away and a fraction of the antibody remains fixed. This fraction is quantitated by incubation with a second antibody–enzyme conjugate followed by addition of enzyme substrate to generate a colour/fluorescence/chemiluminescence that is measured. The amount of DNA adduct in the unknown sample is inversely proportional to this value and is converted to a quantity of adduct by reference to standard curves [26–28].

4.4.3.6. Immunoslotblot assay. DNA is sonicated to give fragments of ~100 base pairs and then heat-denatured. The resulting solution is applied to the filter using a commercially available slot blot appara-

tus so that a small amount of DNA (typically 1–5 μg) is present in each slot. Standards, quality control samples, and unknowns are applied to the same filter in triplicate (typically up to 96 slots per filter). DNA is immobilised on the filter by heat treatment. Filters are then treated with appropriate concentrations of antibody specific for the adduct of interest in the presence of unrelated protein to block non-specific binding. A second antibody–enzyme conjugate is added to the washed filter and allowed to bind. After a further washing step, the filter is soaked in a chemiluminescence reagent and emitted light is detected either directly or by exposure to a suitable photographic film. The intensity of bands corresponding to slots is measured and used to quantitate the level of DNA adduct in unknowns by reference to standard curves generated on the same filter [29,30].

4.4.4. Sources of assay variation

A lack of uniform collection, shipping, processing, and storage of samples will increase technical variability.

4.4.5. Statistical analysis

As described in Section 2, data should be evaluated according to the statistical plan of the study. For each subject, results should be expressed as mole adduct per mol normal base(s). The choice of statistical method to use may depend on the distribution of the data. If the distribution of the data is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, parametric methods based upon normal distribution can also be used, when appropriate, after relevant data transformation. When needed, zero values can be transformed by different ways [31] or by assigning them the value of a fraction of the detection limit. Statistical advice may be needed to determine whether parametric or non-parametric methods are more appropriate for a particular study. The relative and absolute effects of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion by estimating the precision of the effects seen in the study. Stratified analyses can be used for exploring the effect modification by factors such as age, sex, smoking, other exposures, etc. More sophisticated multivariate methods may be necessary for investigating the effect of factors that either may modify or be confounded with the effects. Wherever possible, the analysis should be based on individual exposure information. Consideration for the effect of multiple comparisons on the false positive rate may be included in the statistical analysis.

4.4.6. Interpretation of results

Criteria for identifying a positive response (i.e., increased levels of adducts in an exposed group compared to the referent population; a dose response relationship if variable exposure levels have been established) must be identified prior to implementing the study.

Generally, a statistically significantly higher level of DNA adducts in the exposed population in comparison to the referent population is considered a positive result. A positive result is supported by a dose–response relationship when individuals with different exposure levels are considered. Maximum reliability that an exposure results in increased levels of DNA adducts requires reproducible results in independent studies, such as, analysis by a different method or a similar analysis carried in a different laboratory. It is appreciated that independent studies are not always possible (e.g., in the case of accidents).

The lack of a statistically significant increase in DNA adducts in a particular study indicates, only, that under the exposure conditions evaluated and for the calculated power of the study, the exposure did not result in a increase in DNA damage in the cell population/tissue evaluated.

The results of the quality assurance control are used to demonstrate the adequacy of the methodology for detecting DNA adducts.

4.5. Report

All data should be presented in tabular and/or graphical form, and include all observed replicate results (see below). This will allow comparisons between laboratories, re-analyses of results by outside reviewers, and provide a database for baseline

background values and for exposures to physical and chemical agents. All results of the study should be reported with none omitted, and the definition and handling of ''outliers'' described.

The study report should include the following minimal information:

- copy of the questionnaire used to determine personal characteristics
- statement of ethical approval
- characteristics of the exposed and referent populations
- exposure information and analytical methods used
- tissue collection procedures, including timing in relationship to exposure, transportation, storage, and chain-of-custody
- method for coding samples prior to analysis
- methods for DNA extraction and quantitation
- method to quantitate standard DNA used to generate standard curves
- representative printouts of raw data and standard curves
- for immunochemical methods, the specificity of the primary antibody
- tabulated results for each subject
- statistical methods
- discussion and interpretation of results
- internal and external quality assurance procedures

4.6. Literature cited

1. J.D. Groopman, T.W. Kensler, The light at the end of the tunnel for chemical-specific biomarkers: daylight or headlight? Carcinogenesis 20 (1999) 1–11.
2. K. Hemminki, A. Dipple, D.E.G. Shuker, F.F. Kadlubar, D. Segerbäck, H. Bartsch (Eds.), DNA Adducts, Identification and Biological Significance, International Agency for Research on Cancer, Lyons, 1994, IARC Scientific Publications 125, 478 pp.
3. G.-S. Qian, R.K. Ross, M.C. Yu, J.-M. Yuan, Y.-T. Gao, B.E. Henderson et al., A follow-up study of urinary markers of aflatoxin exposure and liver cancer risk in Shanghai, People's Republic of China, Cancer Epidemiol. Biomarkers Prev. 3 (Suppl. 5) (1994) 3–10.
4. T.W. Kensler, J.D. Groopman, T.R. Sutter, T.J. Curphey, B.D. Roebuck, Development of cancer chemopreventive agents: oltipraz as a paradigm, Chem. Res. Toxicol. 12 (1999) 113–126.
5. S.J. Culp, D.W. Gaylor, W.G. Sheldon, L.S. Goldstein, F.A. Beland, DNA adduct measurement in relationship to small intestinal and forestomach tumour incidence during the chronic feeding of coal tar or benzo(*a*)pyrene to mice. Polycyclic Aromatic Hydrocarbons, 11 (1996) 161–168.
6. L.S. Goldstein, E.H. Wyand, S. Safe, M. Steinberg, S.J. Culp, D.W. Gaylor, F.A. Beland, L.V. Rodriguez, Tumours and DNA adducts in mice exposed to benzo(*a*)pyrene and coal tars: implications for risk assessment. Environ. Health Perspect., 106 (Suppl. 6) (1998) 1325–1330.
7. H. Bartsch, M. Rojas, K. Alexandrov, A.M. Camus, M. Castegnaro, C. Malaveille, S. Anttila, K. Hirvonen, K. Husgafvel-Pursiainen, E. Hietanen, H. Vainio, Metabolic polymorphism affecting DNA-binding and excretion of carcinogens in humans, Pharmacogenetics, 7 (1995) S84–S90.
8. H. Bartsch, E. Hietanen, The role of individual susceptibility in cancer burden related to environmental exposure, Environ. Health Perspect. 104 (1996) 569–577.
9. J. Topinka, B. Binkova, G. Mrackova, Z. Stavkova, V. Peterka, I. Benes, J. Dejmek, J. Lenicek, T. Pilcik, R.J. Sram, Influence of GSTM1 and NAT2 genotypes on placental DNA adducts in an environmentally exposed population, Environ. Mol. Mutagen. 30 (1997) 184–195.
10. D.E.G. Shuker, P.B. Farmer, Relevance of urinary DNA adducts as markers of carcinogen exposure, Chem. Res. Toxicol. 5 (1992) 450–460.
11. C. Zhao, R. Kumar, K. Zahlsen, H. Bager-Sundmark, K. Hemminki, I. Eide, Persistence of 7-(2-hydroxyethyl)guanine-DNA adducts in rats exposed to ethene by inhalation, Biomarkers 2 (1997) 355–359.
12. D. McGregor, D. Anderson, DNA damage and repair in mammalian cells in vitro and in vivo as

indicators of exposure to carcinogens, in: D.B. McGregor, J.M. Rice, S. Venitt (Eds.), IARC Scientific Publications 146. The use of short-ane medium-term tests for carcinogens and data on genetic effects in carcinogenic hazard evaluation. International Agency for Research on Cancer, Lyons, 1999, pp. 309–354.
13. V. Bykov, C.T. Jansen, K. Hemminki, High levels of dipyrimidine dimers are induced in human skin by solar simulating UV-radiation, Cancer Epidemiol. Biomarkers Prev. 7 (1998) 199.
14. S.-M. Hou, K. Yang, F. Nyberg, K. Hemminki, G. Pershagen, B. Lambert, Hprt mutant frequency and aromatic DNA adduct level in nonsmoking lung cancer patients and population controls, Carcinogenesis 20 (1999) 437–444.
15. D.H. Phillips, M. Castegnaro, H. Bartsch (Eds.), Post-labelling methods for detection of DNA adducts. IARC Scientific Publications No. 124 IARC, Lyon, International Agency for Research on Cancer, Lyons, 1993, IARC Scientific Publications 124, 392 pp.
16. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
17. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health, Biosafety in Microbiological and Biomedical Laboratories, HHS Publication No (CDC), 93-8395, 3rd edn. (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H. Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices, 2nd edn., ASM Press, Washington DC, 1995.
18. A.K. Prahalad, J.A. Ross, G.B. Nelson, B.C. Roop, L.C. King, S. Nesnow, M.J. Mass, Dibenzo[*a*,*l*]pyrene-induced DNA adduction, tumorigenicity, and Ki-ras oncogene mutations in strain A/J mouse lung. Carcinogenesis 18 (1997) 1955–1963.
19. W.L. Reichert, J.E. Stein, B. French, P. Goodwin, U. Varanasi, Storage phosphor imaging technique for detection and quantitation of DNA adducts measured by the P-32 postlabeling assay, Carcinogenesis 13 (1992) 1475–1479.
20. D.H. Phillips, Detection of DNA modifications by the postlabelling assay, Mutat. Res. 378 (1997) 1–12.
21. K. Hemminki, H. Rajaniemi, M. Koskinen, J. Hansson, Tamoxifen-induced DNA adducts in leucocytes of breast cancer patients, Carcinogenesis 18 (1997) 9–13.
22. K. Randerath, E. Randerath, G.D. Zhou, D.H. Li, Bulky endogenous DNA modifications (I-compounds) — possible structural origins and functional implications, Mutat. Res. 424 (1999) 183–194.
23. P.B. Farmer, G.M.A. Sweetman, Mass spectrometric detection of carcinogen adducts, J. Mass Spectr. 30 (1995) 1369–1379.
24. G.M.A. Sweetman, D.E.G. Shuker, R.P. Glover, P.B. Farmer, Mass spectrometry in carcinogenesis research, Adv. Mass Spectr. 14 (1998) 343–376.
25. R. Müller, M.F. Rajewsky, Antibodies specific for DNA components structurally modified by chemical carcinogens, J. Cancer Res. Clin. Oncol. 102 (1981) 99–113.
26. J. Adamkiewicz, G. Eberle, N. Huh, P. Nehls, M.F. Rajewsky, Quantitation and visualisation of alkyl deoxynucleosides in the DNA of mammalian cells by monoclonal antibodies, Environ. Health Perspect. 62 (1985) 49–55.
27. C.P. Wild, Antibodies to DNA alkylation adducts as analytical tools in chemical carcinogenesis, Mutat. Res. 233 (1990) 219–233.
28. D.E.G. Shuker, DNA adducts in mammalian cells as indicators of exposure to carcinogens, in: D.B. McGregor, J.M. Rice, S. Venitt (Eds.), IARC Scientific Publication 146 The Use of Short- and Medium-term Tests for Carcinogens and Data on Genetic Effects in Carcinogenic Hazard Evaluation. International Agency for Research on Cancer, Lyons, 1999, pp. 287–308.
29. P. Nehls, J. Adamkiewicz, M.F. Rajewsky, Immuno-slot-blot: a highly sensitive immunoassay for the quantification of carcinogen-modified nucleosides in DNA, J. Cancer Res. Clin. Oncol. 108 (1984) 23–29.
30. C. Leuratti, R. Singh, C. Lagneau, P.B. Farmer, J.P. Plastaras, L.J. Marnett, D.E.G. Shuker, Determination of malodialdehyde-induced DNA damage in human tissues using an immunoslot blot assay, Carcinogenesis 19 (1998) 1919–1924.

31. E.B. Whorton, Some experimental design and analysis considerations for cytogenetic studies, Environ. Mutagen. 7 (Suppl. 4) (1985) 9–15.

5. Single cell gel electrophoresis (COMET) assay

5.1. Introduction

5.1.1. Purpose of the guideline

To provide concise guidance for detecting DNA damage (strand breaks, alkali labile sites [ALS], crosslinking) and incomplete excision repair sites, using the single cell gel (SCG)/Comet assay [1,2], in cells sampled from individuals potentially exposed to genotoxic carcinogens. This guideline must be read in conjunction with the General Introduction (Section 2), which contains information on study design, ethics, and appropriate statistical considerations.

5.1.2. Principles of the assay

Reviews of this assay have been published by McKelvey-Martin et al. [3], Fairbairn et al. [4], Tice [5], Olive et al. [6], Anderson et al. [7], Rojas et al. [8], and Speit and Hartmann [9]. Collins et al. [10] have reviewed specifically the application of this assay to human biomonitoring. There was no attempt to cite all relevant articles in this guideline and, unless otherwise indicated, citations for information provided can be found in these review articles.

Using this technique, virtually any accessible cell population, including but not limited to blood leukocytes (i.e., total, granulocytes, lymphocytes) or bladder, buccal, gastric, nasal, and sperm cells [5,7,8,11–14], can be evaluated for DNA damage. Which cell population is most appropriate will depend on the characteristics of the exposure and the exposed/referent populations, sample timing, the focus of the study, and accessibility.

In this assay, cells suspended in molten agarose are layered onto a microscope slide, the cells lysed by detergents and high salt, and the liberated DNA electrophoresed under neutral or alkaline conditions. Cells with increased levels of DNA damage display altered migration of the DNA toward the anode. Comets (i.e., individual cell DNA migration patterns) are viewed using fluorescence or non-fluorescence microscopy after staining with a suitable dye. Under neutral electrophoretic conditions, DNA migration is increased by double strand breaks (DSB). Under alkaline electrophoretic conditions, it appears that DNA migration is increased preferentially by strand breaks (both single and double) at a pH of 12.1 and by both strand breaks and ALS at higher pH levels [15]. The strand breaks present in DNA prior to alkali treatment arise from direct strand breakage or from incomplete excision repair sites. The expression of ALS as single strand breaks (SSB) increases with increasing pH (maximal expression at $pH > 13$) and increasing exposure duration to alkali prior to electrophoresis. DNA–DNA and DNA–protein crosslinking reduces DNA migration [16,17]. One critically important advantage of this assay is the ability to detect increased levels of damage among subsets of cells within a larger population of apparently unaffected cells.

In addition to the basic assay, the technique has been modified to detect specific classes of DNA adducts (e.g., thymidine dimers, oxidative damage) by using lesion specific antibodies [18] or specific DNA repair enzymes [19–21].

5.1.3. Definitions

Tail length is defined as the maximum length of DNA migration measured either from the estimated leading edge of the head or from the centre of the head, and is considered a measure of the smallest-sized DNA fragments. The percentage of migrated DNA is the fraction of DNA in the tail as compared to the whole image. Tail moment is defined as some measure of tail length multiplied by the fraction of DNA in the tail. There are several methods for calculating tail moment (e.g., Ref. [22,23]) and which method is used is at the discretion of the investigator.

5.1.4. Significance of the endpoint and application in risk assessment

Most human and multi-site, multi-species carcinogens are genotoxic and various types of DNA damage have been characterised for many of them. The relationship between DNA damage, persistence and

repair, and mutagenic endpoints are complex. Because of the demonstrated ability of the alkaline SCG/Comet assay to detect DNA damage in eukaryote cells treated in vitro or in vivo with genotoxic agents and its ease of application, the technique has become increasingly used in human biomonitoring [3–12]. Depending on the electrophoretic condition used (e.g., pH, use of lesion-specific DNA repair enzymes or antibodies), the assay detects single or multiple classes of DNA damage. Unless specific classes of DNA damage identified as being mechanistically involved in the initiation and/or progression of tumourigenicity are evaluated in the relevant cell population, the DNA damage measured identifies hazard rather than risk. The relevance in individuals or groups of positive findings in the SCG/Comet assay to cancer has not been evaluated in a prospective study. Also, due to different mechanisms of formation, the correlation between this assay and other measures of genotoxicity (e.g., chromosomal aberrations, SCEs, mutations) would not be expected to be highly significant for all chemical classes. This means that, depending on the nature of the exposure and the relationship between exposure and sample time, a different biomarker may be preferred.

5.2. Population characteristics

Important characteristics to consider for the exposed and referent populations are discussed in Section 2. As presented in this section, specific information on lifestyle (e.g., smoking, recreational/medicinal drug usage), health status, very recent occurrence of anaerobic exercise [24], and endogenous factors (e.g., age, sex) need to be obtained for each subject. When leukocytes are evaluated for DNA damage, health status is especially important due to the ability of various disease states to activate granulocytes and be accompanied by an increase in DNA damage associated with an oxidative burst. Other disease states associated with free radical generation-induced DNA damage also need to be considered [25,26]. When feasible, genetic polymorphisms associated affecting comet parameters should be identified. A concurrent referent population is required.

Whenever possible, the investigator(s) should have information regarding the types of DNA damage (free radical, SSB, crosslinking) anticipated for the exposure being evaluated. Also, as most environmental and occupational exposures involve complex mixtures, the investigator(s) needs to be aware that the presence of a crosslinking agent may hinder an increase in DNA migration associated with the induction of SSB and/or ALS by other agents in the mixture.

As discussed in Section 2, the number of subjects needed for a particular study design depends, for a desired power, on the magnitude of the expected effect and the known variability of the measured endpoint. Information on optimal population size characteristics for DNA damage in human cells should always be considered during study design. Appropriate software programs are available to assist the investigator in making these calculations. Where the number of subjects to evaluate cannot be determined prior to study implementation or in situations where the appropriate number of subjects are not available, the uncertainty of the response observed should be considered in evaluating the results.

5.3. Sample timing

The optimal sample time for any cell population is during a long-term chronic exposure when the induction and repair of DNA damage is presumed to be at steady-state. This sample time maximises the likelihood of identifying an exposure to a DNA damaging agent. For the sampling of cells after an acute exposure or after termination of a chronic exposure to a genotoxic agent, the optimal collection time for detecting induced DNA damage is most likely within a few hours of exposure termination. The extent of DNA damage in a population of cells decreases with increasing time between exposure termination and sampling due to the loss of DNA damage through DNA repair processes (especially for SSB), the loss of heavily damaged cells through apoptosis or necrosis, and cell turnover. As many tissues are a mixed population of cells with different turnover rates (e.g., granulocytes versus lymphocytes in blood), it is especially important that the kinetics of cell turnover be known for the tissue of interest. However, the rate of the disappearance of the lesions may depend on the agent of exposure. Cells may be collected at a later time but with an increasing

likelihood that a positive genotoxic effect will not be detectable.

5.4. Methods

More detailed description of the SCG/Comet methodology and supporting references can be found in Refs. [3–10].

5.4.1. Quality assurance

A system for recording the collection, transport, and storage of samples should be established to ensure that sample integrity is maintained.

The use of coded samples is critical for eliminating potential bias.

Although seldom employed in past human population studies, it is important to include a quality assurance step to demonstrate the adequacy of the methodology used in the study. Use of subsets of the population (e.g., smokers versus nonsmokers) to retrospectively demonstrate the adequacy of the methodology is unlikely to be generally useful and should be avoided. It would be more appropriate to include, as part of the study design, a quality assurance control among the samples being analysed. This can be accomplished, for example, by including untreated and treated cells with a known level of DNA damage. This approach assesses the reliability of the electrophoretic conditions and the ability of the scorers to appropriately detect DNA damage. This step must be accomplished in such a manner that the identity of the samples remains unknown to the scorer, and that variability in control or treated cells is minimised.

In addition to this internal quality assurance step, inclusion of external quality assurance steps should be considered in the study design. External steps could involve the sharing of replicate samples with an independent laboratory or the independent re-scoring by another laboratory of a certain fraction of the total samples scored.

5.5. Tissue collection

Because of the risk for infectious diseases (e.g., hepatitis, HIV), appropriate precautions must always be followed when handling human tissue samples [27,28]; see Section 2). It is recommended that cell samples be processed as quickly as possible to avoid altered levels of damage associated with storage. Increasing the time between collection and processing may be accompanied by a decline in DNA damage due to DNA repair and/or the loss of heavily damaged cells, or by an increase in DNA damage associated with storage conditions. To minimise the loss of DNA damage due to DNA repair processes, the cell samples should be refrigerated at or below 8°C between collection and processing. It is advisable that the storage conditions be evaluated for possible effects on DNA migration in the target cells being used for biomonitoring. In many situations, sample collection will occur at sites distant from the laboratory and transportation will be necessary. If cells samples are shipped commercially, a minimum–maximum thermometer in the shipping container will provide a measure of the temperature range experienced. If air transportation is used, the cell sample should not be subjected to X-irradiation at security checkpoints. If likely, a sheet of X-ray film can be included in the shipping package to verify whether such an exposure occurred and the information can than be considered during interpretation of the results. If samples are shipped internationally, appropriate permits should be obtained in advance and included in the shipment to avoid delays at customs. Most importantly, samples from both exposed and non-exposed individuals must be handled in the same manner. Where sample collection is protracted across time, a balanced collection design must be used.

Because of its ready accessibility, blood is the tissue used most frequently in biomonitoring studies. If clotting is at all possible after collection, blood samples should be mixed with lithium heparin or sodium heparin. Whole blood leukocytes, subsets of leukocytes (granulocytes, lymphocytes), or even more defined subsets of cells (e.g., naïve, memory T-cells) can be used for an evaluation of DNA damage [5,7,10,29,30]. Total leukocytes may be preferred when exposure conditions are at steady-state or when sampling occurs within a day after exposure. When lymphocytes are being selectively evaluated, the use of unstimulated cells is recommended.

Direct cell samples can be collected by scraping with a spatula, toothbrush, tissue sample brush, or by

other appropriate means, from the desired location (e.g., gastric or nasal mucosa) [11–12]; from urine, cells can be collected by centrifugation. The cells should be collected into culture medium demonstrated to maintain viability. The presence of millimolar concentrations of ethylenediaminetetraacetic acid (EDTA) may be useful for conditions where activation of endonucleases may be anticipated. Any cell population used for a direct assessment of other endpoints can be used for this assay, as long as the potential adverse impact of dying or dead cells on the interpretation of the results is considered.

In studies where the time between sampling and processing is variable and/or excessive, it may be possible or even preferable to evaluate for DNA damage using cryopreserved cells (e.g., Ref. [31]), although differences in migration due to freezing have been reported also. Critical to any sample storage process, is a demonstration that the selected process maintains DNA integrity for the type of damage being assessed (e.g., Ref. [32]).

Selection of which tissue and which cell population to use should be based on the dynamics of the exposure conditions, the persistence of the DNA damage, and the rates of cell turnover. Evaluating both granulocytes and lymphocytes in blood for DNA damage may provide information on recent versus long-term damage, respectively.

5.5.1. Cell viability

The single most important artefact associated with this assay is non-specific DNA damage during cell processing or associated with cell death, mediated either through apoptosis or necrosis. Such processes result in the induction of DNA DSB, which will increase the extent of DNA migration in cell samples. Due to the low exposure levels encountered in most human exposure situations, the frequency of dying/dead cells in blood should be less than a few percent and are, therefore, probably not an issue. In contrast, dying/dead cells should be common in terminally differentiated cell populations with a high renewal rate, such as buccal or bladder cells. Several dye-based assays, such as 5-6 carboxyfluorescein diacetate/ethidium bromide [33] have been used to assess viability. Another method would be to assess for the presence of cells with very low molecular weight DNA indicative of apoptosis or necrosis [34]. When increased levels of DNA migration are present in cells sampled from the exposed population, the possibility that the increase is due to dying/dead cells needs to be formally evaluated.

5.5.2. Slide preparation and cell lysis

Slides are made using a suitable technique that results in an appropriate distribution of cells throughout the gel such that scoring bias due to excessive cell density is avoided [9]. Regardless of the methodology used, it is highly recommended that the one used be suitable for the preparation of dried gels after neutralisation [35]. Dried slides enable not only the ability to score slides when convenient after electrophoresis but also allow the slides to be archived/stored under suitable conditions for reanalysis, if deemed necessary. Cell lysis, to remove lipids and proteins, is accomplished using standard procedures suitable for the cell type being evaluated. If appropriate, residual protein can be removed using proteinase K.

5.5.3. Detection of specific classes of DNA damage using DNA from lysed cells

After lysis, the DNA of individual cells in the gel can be incubated with lesion-specific DNA repair enzymes (e.g., endonuclease III for oxidised pyrimidines, formamido pyrimidine glycosylase for 8-hydroxyguanine lesions) to express these lesions as single strand breaks [10].

5.5.4. Electrophoresis and staining

After removal from the lysing solution, the slides are drained or rinsed and placed in a horizontal gel electrophoresis tank for DNA unwinding (if under alkaline conditions) and electrophoresis. The unwinding/electrophoresis conditions used should be based on the types of damage being assessed and on the extent of DNA migration associated with the conditions and cell/tissue type of interest. In studies where specific classes of DNA adducts or cross-linking are not the focus, alkaline unwinding conditions that maximise the expression of ALS as DNA damage should be used. Generally, this means using a pH > 13 and unwinding conditions as long as 60 min (or more) [36]. In contrast, in studies using repair enzymes to cleave sites where specific classes of DNA damage are present, a pH more specific for

SSB only (i.e., 12.1) may be more suitable [15]. The subsequent electrophoretic conditions (i.e., voltage, amperage, duration) can be varied to achieve the desired extent of migration for representative ''control'' cells. In situations where crosslinking is a possibility, it is important to have an appreciable level of DNA migration in the referent cell population, such that a decrease in DNA migration would be evident. Once optimised, the same electrophoretic conditions must be used throughout the study. All studies should be conducted in a manner that minimises or eliminates intra- and inter-run differences in DNA migration.

The electrophoresis buffer is neutralised and the slides are either dried prior to staining/scoring [35] or, immediately after neutralisation, the DNA is stained with an appropriate fluorescent (e.g., ethidium bromide, propidium iodide, 4′,6-diamidino-2-phenylindole) or nonfluorescent (e.g., silver nitrate) dye [37]. The DNA can also be stained using antibodies specific for certain classes of DNA damage [18].

5.5.5. Slide scoring

Slides must be scored without knowledge of the exposure group; slides from the quality assurance standard should be included within this set of slides. Where possible, the number of cells to be analysed per subject should be based on relevant historical data. As this is not always possible, at least a 100 cells per subject should be scored (50 per each of two replicate slides is recommended to reduce possible position effects during electrophoresis). This minimal number of cells is based on a desire to detect the presence of small subpopulations of cells with increased damage in an otherwise unaffected population of cells. Optimal magnification may vary depending on the type of cell being scored but is generally $200\times$ to $400\times$. DNA migration can be assessed using a variety of methods, including but not limited to visual (e.g., using a micrometer to estimate image length or tail length or using a qualitative estimate of the amount of migrated DNA in arbitrary units; photography followed by using a ruler or a grid; or using a computer-based image analysis system to measure such parameters as tail length, the percentage of migrated DNA, and tail moment) [3–10].

5.5.6. Sources of assay variation

A lack of uniform collection, shipping, processing, and storage of samples will increase technical variability.

A lack of uniformity in regard to reagents/solutions, preparation of agarose gel slides, and conditions for lysis, alkali unwinding, and electrophoresis will increase technical variability.

Scoring of comets is a potential source of variability. Variation can be minimised by using only one well-trained scorer, and, when several scorers are used, reduced by ensuring a balanced scoring design.

Also, when more than one scorer is used, it is advisable that possible scorer differences be evaluated by cross-scoring a set of slides.

5.5.7. Statistical analysis

As described in Section 2, data should be evaluated according to the statistical plan of the study.

For each subject, for the measure of DNA migration used, the mean (plus standard deviation or standard error of the mean), median (plus confidence interval), range, n (i.e., the number of cells), and, if possible, some measure of the distribution of migration among cells should be presented. In studies where the variate of interest is the tail moment, individual response data on the percentage of migrated DNA and tail length should also be provided in order to evaluate for concordance among the two endpoints.

There is no unified agreement on the proper statistical methods to use [38]. However, the unit of exposure is the individual and not the cell; pooling cells across individuals within a study group eliminates a critically important source of variability. The choice of statistical method to use may depend on the distribution of the comet parameter data. If the distribution of comet parameters is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, parametric methods based upon normal distribution can also be used, when appropriate, such as after relevant data transformation. Statistical advice may be needed to determine whether parametric or non-parametric methods are more appropriate for a particular study. In addition, an analysis based on the distribution of comets within each individual could

be conducted. One possible approach is to analyse the effect of the exposure on the dispersion coefficient H, where H = variance/mean [31,39]. In the estimation approach, the relative and absolute effects of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion derived from the hypothesis testing by estimating the precision of the effects seen in the study. Stratified analyses can be used for exploring the effect modification by factors such as age, sex, smoking, other exposures, etc. More sophisticated multivariate methods may be necessary for investigating the effect of factors that either may modify or be confounded with the effects. Wherever possible, the analysis should be based on individual exposure information. Consideration for the effect of multiple comparisons on the false positive rate may be included in the statistical analysis.

The quality assurance standard is compared against the relevant concurrent control samples using appropriate statistics.

5.5.8. Interpretation of results

Criteria for identifying a positive response (i.e., increased levels of DNA damage in an exposed group compared to the referent population; a dose–response relationship if variable exposure levels have been established) must be identified prior to implementing the study.

Generally, a statistically significant increase or decrease in DNA migration in the exposed population in comparison to the referent population is considered a positive result. Presentation of the results in terms of the size of the effect with confidence intervals may also help with the interpretation. A positive result is supported by a dose–response relationship when individuals with different exposure levels are considered. Maximum reliability that an exposure results in increased levels of DNA damage requires reproducible results in independent studies. It is appreciated that independent studies are not always possible (e.g., in the case of accidents).

The lack of a statistically significant increase in DNA damage in a particular study indicates, only, that under the exposure conditions evaluated and for the calculated power of the study, the exposure did not result in a increase in DNA damage in the cell population/tissue evaluated.

The results of the quality assurance control are used to demonstrate the adequacy of the SCG/Comet methodology for detecting DNA damage.

5.6. Report

All data should be presented in tabular form, and should include all observed results. This will allow comparisons between laboratories, re-analyses of results by outside reviewers, and provide a database for baseline background values and for exposures to physical and chemical agents. All results of the study should be reported with none omitted, and the definition and handling of ‘‘outliers’’ described.

At a minimum, the study report should include the following information:

- copy of the questionnaire used to determine personal characteristics
- statement of ethical approval
- characteristics of the exposed and referent populations
- exposure information and analytical methods used
- tissue collection procedures, including timing in relationship to exposure, transportation, storage, and chain-of-custody
- method for coding samples prior to analysis
- methods for sample processing, agarose gel slide preparation, lysis, alkali unwinding, electrophoresis, neutralisation, fixation/drying, staining, and scoring
- tabulated results for each subject
- statistical methods
- discussion and interpretation of results
- internal and external quality assurance procedures.

5.7. Literature cited

1. O. Östling, K.J. Johanson, Microelectrophoretic study of radiation-induced DNA damages in individual mammalian cells, Biochem. Biophys. Res. Commun. 123 (1984) 291–298.
2. N.P. Singh, M.T. McCoy, R.R. Tice, E.L. Schneider, A simple technique for quantitation of low levels of DNA damage in individual cells, Exp. Cell Res. 175 (1988) 184–191.

3. V.J. McKelvey-Martin, M.H.L. Green, P. Schmezer, B.L. Pool-Zobel, M.P. De Méo, A. Colins, The single cell gel electrophoresis assay (Comet assay): a European review, Mutat. Res. 288 (1993) 47–63.
4. D.W. Fairbairn, P.L. Olive, K.L. O'Neill, The Comet assay: a comprehensive review, Mutat. Res. 339 (1995) 37–59.
5. R.R. Tice, The single cell gel/comet assay: a microgel electrophoretic technique for the detection of DNA damage and repair in individual cells, in: D.H. Phillips, S. Venitt (Eds.), Environmental Mutagenesis Bios Scientific Publishers, Oxford, 1995, pp. 315–339.
6. P.L. Olive, P.J. Johnston, J.P. Banath, R.E. Durand, The comet assay: a new method to examine heterogeneity associated with solid tumors, Nat. Med. 4 (1998) 103–105.
7. D. Anderson, T.-W. Yu, D.B. McGregor, Comet assay responses as indicators of carcinogen exposure, Mutagenesis 13 (1998) 539–555.
8. E. Rojas, M.C. Lopez, M. Valverde, Single cell gel electrophoresis assay: Methodology and applications, J. Chromatogr. B: Biomed. Appl. 722 (1999) 225–254.
9. G. Speit, A. Hartmann, The Comet assay (single-cell gel test) — A sensitive genotoxicity test for the detection of DNA damage and repair, DNA Repair Protoc. 113 (1999) 203–212.
10. A. Collins, M. Dusinska, M. Franklin, M. Somorovska, H. Petrovska, S. Duthie et al., Comet assay in human biomonitoring studies: Reliability, validation, and applications, Environ. Mol. Mutagen. 30 (1997) 139–146.
11. B.L. Pool-Zobel, N. Lotzmann, M. Knoll, F. Kuchenmeister, R. Lambertz, U. Leucht et al., Detection of genotoxic effects in human gastric and nasal mucosa cells isolated from biopsy samples, Environ. Mol. Mutagen. 24 (1994) 23–45.
12. L. Calderón-Garciduenas, N. Osnaya-Brizuela, L. Ramirez-Martinez, A. Villarreal-Calderon, DNA strand breaks in human nasal respiratory epithelium are induced upon exposure to urban pollution, Environ. Health Perspect. 104 (1996) 145–154.
13. C.M. Hughes, V.J. McKelvey Martin, S.E.M. Lewis, Human sperm DNA integrity assessed by the comet and ELISA assays, Mutagenesis 14 (1999) 71–75.
14. D. Anderson, M.M. Dobrzynska, N. Basaran, Effect of various genotoxins and reproductive toxins in human lymphocytes and sperm in the comet assay, Teratogen. Carcinogen. Mutagen. 17 (1997) 29–43.
15. Y. Miyamae, K. Iwaski, N. Kinae, S. Tsuda, M. Murakama, M. Tanaka, Y.F. Sasaki, Detection of DNA lesions induced by chemical mutagens by the single cell gel electrophoresis (comet) assay. II. Relationship between DNA migration and alkaline conditions, Mutat. Res. 393 (1997) 107–113.
16. S. Pfuhler, H.U. Wolf, Detection of DNA-crosslinking agents with the alkaline comet assay, Environ. Mol. Mutagen. 27 (1996) 196–201.
17. O. Merk, G. Speit, Detection of crosslinks with the comet assay in relationship to genotoxicity and cytotoxicity, Environ. Mol. Mutagen. 33 (1999) 167–172.
18. S. Sauvaigo, C. Serres, N. Signorini, N. Emonet, M.J. Richard, J. Cadet, Use of the single-cell gel electrophoresis assay for the immunofluorescent detection of specific DNA damage, Anal. Biochem. 259 (1998) 1–7.
19. A.R. Collins, S.J. Duthie, V.L. Dobson, Direct enzymatic detection of endogenous base damage in human lymphocyte DNA, Carcinogenesis 14 (1993) 1733–1735.
20. A.R. Collins, M. Dusinska, C.M. Gedik, R. Stetina, Oxidative damage to DNA: do we have a reliable biomarker? Environ. Health Perspect. 104 (1996) 465–469.
21. G. Speit, S. Haupter, P. Schütz, P. Kreis, Comparative evaluation of the genotoxic properties of potassium bromate and potassium superoxide in V79 Chinese hamster cells, Mutat. Res. 439 (1999) 213–221.
22. P.L. Olive, J.P. Banath, Induction and rejoining of radiation induced DNA single-strand breaks: ''tail moment'' as a function of position in the cell cycle, Mutat. Res. 294 (1993) 275–283.
23. B. Hellman, H. Vaghef, B. Boström, The concepts of tail moment and tail inertia in the single cell gel electrophoresis assay, Mutat. Res. 336 (1995) 123–131.
24. A. Hartmann, U. Plappert, K. Raddatz, M.

Grünert-Fuchs, G. Speit, Does physical activity induce DNA damage? Mutagenesis 9 (1994) 269–272.
25. A.R. Collins, K. Raslova, M. Somorovska, H. Petrovska, A. Ondrusova, B. Vohnout, R. Fabry, M. Dusinska, DNA damage in diabetes: correlation with a clinical marker, Free Radical Biol. Med. 25 (1998) 373–377.
26. M. Hannon-Fletcher, C. Hughes, K.W. Moles, C.R. Barnett, Y.A. Barnett, An investigation of in vivo antioxidant status and DNA damage in patients with IDDM, in: T.K. Basu, N.J. Temple, M.L. Garg (Eds.), Antioxidants in Human Health, CABI Publishing, Oxon, UK, 1999, pp. 259–269.
27. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
28. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health. Biosafety in Microbiological and Biomedical Laboratories. HHS Publication No (CDC) 93-8395, 3rd edn. (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H. Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices, 2nd edn., ASM Press, Washington DC, 1995.
29. G.H.S. Strauss, W.P. Peters, R.P. Everson, Measuring DNA damage in individual cells of heterogeneous mixtures: a novel application of an immunological typing technique, Mutat. Res. 304 (1994) 211–216.
30. R.R. Tice, G.H.S. Strauss, The single cell gel electrophoresis/comet assay: a potential tool for detecting radiation-induced DNA damage in humans, in: T.M. Fliedner, E.P. Cronkite, V.P. Bond (Eds.), Assessment of Radiation Effects by Molecular and Cellular Approaches, V. Stem Cells 13, S1, AlphaMed Press, Dayton, OH, 1995, pp. 207–214.
31. R.R. Tice, G.H.S. Strauss, W.P. Peters, High-dose combination alkylating agents with autologous bone marrow support in patients with breast cancer: preliminary assessment of DNA damage in individual peripheral blood lymphocytes using the single cell gel electrophoresis assay, Mutat. Res. 271 (1992) 101–113.
32. A. Gurugunta, M. Vasquez, R.R. Tice, Effect of different cryopreservation methods on DNA migration in the Single Cell Gel (SCG) assay, Environ. Mol. Mutagen. 27, S27 (1996) 28.
33. G.H.S. Strauss, Non-random cell killing in cryopreservation: implications for performance of the battery of leukocyte tests (BLT) I. Toxic and immunotoxic effects, Mutat. Res. 252 (1991) 1–15.
34. R.R. Tice, M. Furedi-Machacek, D. Satterfield, A. Udumudi, M. Vasquez, J.K. Dunnick, Measurement of micronucleated erythrocytes and DNA damage during chronic ingestion of phenolphthalein in transgenic female mice heterozygous for the *p53* gene, Environ. Mol. Mutagen. 31 (1998) 113–124.
35. M. Klaude, S. Eriksson, J. Nygren, G. Ahnström, The Comet assay: mechanisms and technical considerations, Mutat. Res. 363 (1996) 89–96.
36. Vijayalaxmi, R.R. Tice, G.H.S. Strauss, Assessment of radiation-induced DNA damage in human blood lymphocytes using the single cell gel electrophoresis technique, Mutat. Res. 271 (1992) 243–252.
37. N. Kizilian, R.C. Wilkins, P. Reinhardt, C. Ferrarotto, J.R.N. McLean, J.P. McNamee, Silver-stained comet assay for detection of apoptosis, Biotechniques 27 (1999) 926–930.
38. D.P. Lovell, G. Thomas, R. Dubow, Issues related to the experimental design and subsequent statistical analysis of in vivo and in vitro Comet studies, Teratogen. Carcinogen. Mutagen. 19 (1999) 109–119.
39. B. Margolin, M.D. Shelby, Sister chromatid exchanges: a re-examination of the evidence for sex and race differences in humans, Environ. Mutagen. 7 (1985) 63–72.

6. Mutations of the *HPRT* gene in T-lymphocytes

6.1. Introduction

6.1.1. Purpose of the guideline

To provide concise guidance for the performance and interpretation of assays that measure the fre-

quencies of *HPRT* deficient T-lymphocytes, identified as purine analogue resistant cells, in the blood of individuals potentially exposed to mutagenic agents. Such cells arise by somatic mutation of the *HPRT* gene. This guideline must be read in conjunction with the General Introduction (Section 2), which contains information on such issues as study design, ethics, and appropriate statistical considerations.

6.1.2. Principles of the assays

The *HPRT* gene controls the enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT) which catalyses the reaction of hypoxanthine and guanine with phosphoribosylpyrophosphate for purine salvage. In addition to its normal substrates, HPRT catalyses the conversion of purine analogues such as 6-thioguanine (6-TG), rendering them cytotoxic to normal cells. Following mutation of the gene, cells deficient in HPRT survive treatment with 6-TG as they cannot phosphoribosylate the analogue. The HPRT assays are positive selection assays for in vivo somatic gene mutations [1,2].

HPRT inactivating mutations occurring either in bone marrow stem cells or in differentiated T-cells in peripheral lymphoid pools contribute to the overall frequency of HPRT deficient T-lymphocytes in blood. However, as there is little contribution from the bone marrow stem cells to peripheral T-cell pools after adolescence, the origin of HPRT deficient cells in the blood is, in part, a function of age. In the foetus and in young children, HPRT mutant T-cells may originate from stem, ''pre-thymic'', or differentiating T-cells, as well as from fully differentiated cells [3,4]. By contrast, in adults the predominant cells of origin are the differentiated functionally competent T-lymphocytes that undergo periodic amplification and regression cycles in the periphery, a capacity that also may vary with age [2]. These T-cell kinetics contribute to potentially large variations in the background frequencies of HPRT deficient cells in normal individuals [2,5].

6.1.2.1. Definitions. The term ''HPRT variant'' refers to 6-TG resistant T-cells that are identified on the phenotypic level only (i.e., as cells capable of synthesising DNA in vitro after proper stimulation in the presence of selection). As these variants are detected only by a method that kills them (e.g., autoradiography), the mutational basis of the phenotypic change is presumed but cannot be verified. In contrast, the term ''HPRT mutant'' refers to 6-TG resistant T-cells that are identified by a cloning assay that requires cell proliferation in vitro with growth of a visible colony from which DNA can be extracted and the mutational alteration demonstrated. In practice, as DNA alterations have been routinely demonstrated in T-cell colonies growing in TG, such colonies are usually considered to be mutants without further DNA analysis. A variant frequency (VF) is a measure of HPRT variants (presumed mutations) in peripheral blood as determined by autoradiography or similar assay; a mutant frequency (MF) is a measure of HPRT mutants (verifiable mutations) in peripheral blood as determined by cloning assay. At the extremes, inter-individual VF variability has ranged from 10- to 30-fold in reported studies, with repeat sampling from individual subjects showing up to four-fold differences [2]. Again at the extremes, inter-individual MF variability has shown approximately a 100-fold range for all studies performed in four different laboratories, with repeat sampling from individual subjects showing up to eight-fold differences [2,5]. The greater variability seen in cloning assay results may be due to the ability of that assay to recognise in vivo expansions of mutant clones. However, despite this variability among and within individuals, group mean VFs and MFs for normal populations are usually within narrow ranges and similar in different laboratories [5].

6.1.2.2. Significance of the endpoints and application in risk assessment. The *HPRT* gene is capable of reflecting a wide variety of genetic alterations such as DNA base pair substitutions, large and small deletions, inversions and heterologous chromosome recombinations [1,2]. However, as the gene is located on the X-chromosome, mutations here cannot reflect homologous chromosome recombination resulting in a loss of heterozygosity.

An increase in HPRT deficient T-cells in an exposed population relative to a suitable reference population demonstrates a genotoxic effect in the former. Under appropriate study conditions, an increase in the mean (or median) VF or MF of an exposed population may be taken as a measure of the exposure. However, because of inter-individual vari-

ability in baseline VFs or MFs, exposures so determined are group rather than individual attributes. Direct comparisons of *HPRT* mutations to other endpoints such as haemoglobin adducts, SCEs or chromosome aberrations for detecting known chemical exposures have shown mutations to be generally the least sensitive [6]. However, population exposures are often to unspecified chemicals. Furthermore, *HPRT* mutations have detected even low dose subacute occupational radiation exposures several years after the event [7]. Although the implications of elevated VFs or MFs for cancer risk have not been assessed in prospective human studies, molecular analyses of in vivo derived *HPRT* mutations have shown large deletions with breakpoints at topoisomerase II cleavage consensus sequences, the formation of fusion genes, and V(D)J recombinase-mediated deletions similar to mutagenic changes seen in cancer-related genes or genomic regions associated with cancer [1,2]. Therefore, such cancer related mutational mechanisms are reported by the *HPRT* gene. Furthermore, in vivo T-cell *HPRT* mutations in animals can be induced by known carcinogens [8–10].

6.2. Population characteristics

Important characteristics to consider for the exposed and referent populations are discussed in Section 2. As presented in this section, specific information on lifestyle (e.g., smoking, alcohol consumption, recreational/medical drug usage), health status (e.g., immunological status), and endogenous factors, such as age and sex, need to be obtained for each subject. Confirmation of smoking status by cotinine or similar analysis could be useful. Immunological status, as defined here, refers to the presence of deficiency, hypersensitivity, or autoimmune diseases as well as recent viral infections or immunisations. Low serum folate has been associated with elevated MFs in chemotherapy treated patients, making it important that individuals with dietary or other obvious reasons for low serum folate not be included in either the exposed or the referent population [11]. Study subjects, in general, should not be taking any mutagenic drugs or medications, or agents that select in vivo for or against HPRT mutants. Specific agents such as azathioprine, 6-mercaptopurine and 6-TG select in vivo for HPRT deficient cells, thereby enormously elevating their frequency, while others, such as folate antagonists, could select against and thereby potentially lower their frequency. Wherever possible, individuals with genetic polymorphisms associated with increased susceptibility to genotoxic agents should be identified, as individuals with known genetic instability syndromes have been shown to have increased frequencies of *HPRT* mutations [1,2]. A concurrent referent population is required for all monitoring studies.

As discussed in Section 2, the number of subjects needed for a particular study design depends, for a desired power, on the magnitude of the expected effect and the known variability of the measured endpoint. Because of the high intrinsic inter-individual variability in the frequency of HPRT deficient T-cells in peripheral blood, it is essential that a sufficiently large number of subjects are examined in both the exposed and the referent populations to document any increase in HPRT deficient cells in the former. Appropriate software programs are available to assist the investigator in making these calculations (see Section 2.1). Information on population size characteristics for this endpoint can be found in Robinson et al. [5].

6.3. Sample timing

The optimal sample time is during a long-term chronic exposure as long as the exposure conditions are relatively stable. For peak exposures or for single acute exposures, it will take at least 2 weeks for *HPRT* variants and 2 months for HPRT mutants to be optimally detected in peripheral blood [2,12]. As assays performed prior to these times may underestimate exposure-related VFs or MFs, the timing of peak or single acute exposures should be known. The half-lives of variants or mutants in peripheral blood are unknown and likely variable.

6.4. Methods

6.4.1. Quality assurance

A system for recording the collection, transport, and storage of samples should be established to

ensure that sample integrity is maintained. The use of coded samples is critical for eliminating potential bias. Blinding and coding usually occur at the blood collection stage. Although seldom employed in past human population studies, it is important to include a quality assurance step to demonstrate the adequacy of the methodology used in the study. In some situations, subsets of the population (e.g., smokers versus nonsmokers, aged versus young) may serve, retrospectively, to demonstrate the adequacy of the methodology. However, this approach is unlikely to be generally useful and should be avoided. It would be more appropriate to include, as part of the study design, a quality assurance control among the samples being analysed. An ideal standard would be a cryopreserved blood sample from a single donor for which the VF or MF is known. Re-assay of this standard with each study allows a precise specification of current laboratory performance. It is imperative that the quality assurance sample be processed using the same methodology as that used in the population study and in such a manner that it cannot be uniquely identified. In addition to these internal quality assurance steps, inclusion of external quality assurance steps should be considered in the study design. External steps could involve the sharing of replicate blood samples with an independent laboratory or the independent re-analysis by another laboratory of a certain fraction of the total samples analysed. Although absolute VF or MF values will probably depend on specific laboratory protocols [13], relative differences between exposed and referent groups should be similar in different laboratories.

6.4.2. Blood collection

Because of the risk for infectious diseases (hepatitis, HIV), appropriate precautions must always be followed when handling blood samples ([14,15]; see Section 2.1) At the time of collection, blood is mixed with heparin to prevent coagulation. It is recommended that blood samples be processed within 24 h of collection. Processing, as used here, means isolation of mononuclear cells and cryopreservation in liquid nitrogen, or the immediate assay of freshly isolated lymphocytes, as may be done for the cloning assay. If longer intervals elapse before processing, the interval must be recorded. In many instances, blood collection will occur at sites distant from the laboratory and transportation will be necessary. If fresh anticoagulated blood samples are transported commercially, the range of temperature variations should be recorded, as should the condition of the sample at receipt. Blood samples should preferentially be maintained at room temperature. If air transportation is used, the blood should not be subjected to X-irradiation at security checkpoints. If possible, a sheet of X-ray film can be included in the shipping package to verify that such an exposure occurred and the information considered during interpretation of the results. If samples are shipped internationally, appropriate permits should be obtained in advance and included in the shipment to avoid delays at customs. Blood from both exposed and referent individuals must be handled in the same manner and concurrently. Where blood collection is protracted across time, a balanced collection design should be used. An alternative method for sending mononuclear cell (MNC) samples over long distances is to initially separate and cryopreserve this cell fraction at a local laboratory and then send the frozen cells in liquid nitrogen ''dry shippers'' by air freight. Dry shippers are accepted for transport by the airlines.

6.4.3. Description

Two methods have been developed for measuring purine analogue resistant lymphocytes in peripheral blood. These are: (i) a short term method termed the autoradiographic assay, used to measure variants, as defined above, and (ii) a longer term method that requires cell growth, termed the cloning assay, to measure mutants, as defined above. There are large methodological differences between the two methods, which may lead to somewhat different results. For either assay, the MNC fraction is separated from heparinised peripheral blood by density centrifugation, using a Ficoll/sodium metrizoate or similar preparation. After washing cells in phosphate-buffered saline and counting for viability, the cells may be used immediately for the cloning assay or cryopreserved in liquid nitrogen either for the autoradiographic assay or the cloning assay, using culture medium supplemented with serum and dimethylsulfoxide (DMSO). Details of the two assays are described separately. The critical step of cryopreservation, which is mandatory for the autoradiographic

assay, but optional for the cloning assay, involves controlled freezing under described methods [16,17].

6.4.4. Assay methods

6.4.4.1. Autoradiographic assay [18,19]. MNCs are carefully thawed and stimulated with phytohaemagglutinin (PHA) in replicate short-term cultures, with or without 6-thioguanine (6-TG) (2×10^{-4} M), and incubated until culture termination at 42 h. Cryopreservation is required to avoid labeling of normal T-lymphocytes that are ''in cycle'' in vivo. Typically, tritiated thymidine is added to cultures during the last 18 h to label cells synthesising DNA. At culture termination, cells are treated to obtain free nuclei, fixed, counted and added in measured volumes to microscope slides, and then stained, autoradiographed, and scored blind by microscopy. Bromodeoxyuridine can be used in the place of tritiated thymidine [20–23]. A count is made of all of the rare, labelled nuclei on slides prepared with cells from the TG-containing selection cultures. For the slide prepared from the non-selection labeling index culture, a random differential count is made of a total of 2500 to 3000 labelled and unlabelled cells to identify the Labeling Index (LI). The VF is calculated by dividing the total number of labelled nuclei derived from the TG-containing cultures by the number of evaluatable cells (the number of nuclei recovered from the selection cultures multiplied by the LI). An optimum assay should have high recovery and viability ($> 80\%$) after careful thawing of cryopreserved cells and, after assaying and scoring, should include at least one million evaluatable cells. Lower numbers of evaluatable cells may still allow an acceptable assay, having however, a larger confidence interval around the observed VF.

6.4.4.2. Cloning assay [16,24]. Either fresh or carefully thawed cryopreserved MNCs are counted and primed by seeding approximately 10^6 cells/ml in culture medium containing PHA and incubated at 37°C for 24–40 h. An alternative method allows for omitting the separate priming step, and inoculating MNCs directly into wells where effective priming occurs in the microtitre wells. After priming, if this is done as a separate step, or directly, if it is not, the cells are counted and plated into microtitre wells of 96-well microtitre plates at low cell numbers (usually 1 to 10 cells/well) in non 6-TG containing medium (non-selection) for determining the non-selection cloning efficiency (CE), and at 2×10^4 cells/well in 10^{-5} M 6-TG (selection) for determining the selection CE. Culture medium is supplemented with a suitable source of serum pretested for T-cell cloning and human interleukin-2 (IL-2). All cultures contain, in addition, appropriate lethally irradiated human HPRT deficient lymphoblastoid cells such as TK6. These accessory cells are necessary for optimum T-cell growth in microtitre plates, but it is not known if this is due to metabolic factors or to the provision of a second immunological signal. Plates are incubated to allow a sufficient time period for T-cell colony growth (12 to 17 days) in positive wells, after which the wells are scored for positive and negative wells (no growing colonies) using an inverted microscope. Since the distribution of growing clones among wells follows a Poisson distribution, the CE for both non-selection and selection wells is calculated as: $CE = -\ln P_0/n$, where P_0 is the proportion of negative wells in the non-selection or selection plates and n is the number of cells plated/well in these plates. As in the autoradiographic assay, an optimum cloning assay should have high viability in MNCs prior to use ($> 80\%$) and score at least 10^6 evaluatable cells (non-selection CE $\times$ total cells scored in all selection plates). Again, lower numbers of evaluatable cells will result in higher confidence intervals around the observed MFs.

A MF (per 10^{-6} cells) is then calculated as:

$$MF = CE(\text{selection})/CE(\text{non-selection})$$

The cloning assay allows for the isolation and further propagation of wild-type and mutant T-cell colonies for molecular analysis of the *HPRT* and T-cell receptor (TCR) genes. The former permits an analysis of mutational spectra, which may be of value for recognising specificity of mutation induction by different physical and chemical agents. The latter provides a molecular signature of in vivo clonality for HPRT mutants. These are not considered further here except to indicate that extensive in vivo clonality of one or more T-cell mutations in an individual outlier may be a sufficient reason to eliminate the respective data point(s) in some statistical

analyses of group differences. Other than eliminating outliers, it is not necessary to ''correct'' MFs for clonally amplified single mutational events of the degree usually observed in normal subjects [25].

6.4.5. Sources of assay variation

6.4.5.1. Autoradiographic assay. T-lymphocytes that are cycling in vivo may become labelled in vitro under conditions of the assay unless rigorous methods such as cryopreservation are taken to completely eliminate them from being scored. Without elimination, many T-cells that are in cell cycle in vivo cannot be stopped from entering the first round of DNA synthesis in vitro even in the presence of 6-TG. Cryopreservation appears to arrest such cells at the $G_{0/S}$ checkpoint of the cell cycle [26]. When then placed in culture, these cells proceed into S-phase almost immediately and are no longer in DNA synthesis when tritiated thymidine is added at 24 h after mitogen-stimulation. They do not therefore become labelled and are not scored. As only rare labelled cells are scored, Poisson errors in counting small numbers can be large. Observer bias may affect decisions on rare labelled cells.

6.4.5.2. Cloning assay. As all investigators have reported a strong inverse correlation between non-selection CEs and calculated MFs, assays with low non-selection CEs should be treated with caution. Observer error with failure to recognise small mutant colonies results in underscoring.

6.4.6. Statistical analysis

As described in Section 2.1, VF or MF data should be evaluated according to the statistical plan of the study. The choice of statistical method to use may depend on the distribution of the data. If the distribution of the data is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, parametric methods based upon normal distribution can also be used, when appropriate, after relevant data transformation. Statistical advice may be needed to determine whether parametric or non-parametric methods are more appropriate for a particular study. The relative and absolute effect of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion by estimating the uncertainty of the effects seen in the study. Stratified analyses can be used for exploring the effect modification by factors such as age, sex, smoking, other exposures, etc. More sophisticated multivariate methods may be necessary for investigating the effect of factors that either may modify or be confounded with the effects.

6.4.7. Interpretation of results

Interpretation of results is complicated by inter-individual variation and even the intra-individual variations with time that have been reported by different laboratories. The exposed and referent populations may have VF or MF values within the same range with only the mean or median differing between the two groups. The independent effects of an exposure on non-selection CE values in a cloning assay study may further complicate interpretation.

A statistically significantly higher VF or MF among the exposed population in comparison to an appropriate referent population is considered a positive result (i.e., the agent of interest induced gene mutations in vivo in lymphocytes of exposed individuals). A positive result is supported by a dose response relationship when individuals with different exposure levels are considered. Maximum reliability that the exposure results in an increased VF or MF requires reproducible results in independent studies. It is appreciated that independent assessments are not always possible (e.g., in the case of accidents).

When the cloning assay is performed, information on the types of mutations induced (mutation spectrum) may aid in identifying the nature of the mutagenic agent (e.g., point mutations, deletions, etc.) and increase the specificity and, thereby, sensitivity of the assay.

The lack of a statistically significant increase in VFs or MFs in the exposed population indicates that under the exposure conditions evaluated and for the calculated power of the study, the exposure conditions did not result in a significant increase in gene mutations in vivo.

VF values usually show positive associations with donor age; MF values usually show positive associations with donor age and, in most laboratories, negative associations with non-selection CE values.

Methods by which this is incorporated into a laboratory's analysis of results should be given.

The results of the quality assurance control demonstrate the adequacy of laboratory performance for the study.

6.5. Report

All data should be presented in tabular form, and include all observed results (see below). This will allow comparisons between laboratories, re-analyses of results by outside reviewers, and will provide a database for baseline background values and for exposures to physical and chemical agents. All results of the study should be reported with none omitted, and the definition and handling of ''outliers'' described. Extensive in vivo clonality of *HPRT* mutations in an individual, determined by molecular analysis as described above, may identify the individual as an outlier.

The study report should include the following minimal information:

- copy of the questionnaire used to determine personal characteristics
- statement of ethical approval
- characteristics of the exposed and referent populations
- exposure information and analytical methods used
- blood collection procedures, including timing in relationship to exposure, transportation, storage, and chain-of-custody
- method for coding samples prior to analysis
- the amount of blood obtained
- the method for MNC isolation
- MNC counts and viability after MNC isolation
- the method of MNC storage (i.e., the conditions of cryopreservation)
- cell counts and viability after thawing of cryopreserved cells and after primings for the cloning assay
- the assay conditions (i.e., composition of media, sources of sera, IL-2 and 6-TG, identification and treatment of feeder cells, 6-TG concentration, incubation conditions, time of scoring)
- for the autoradiographic assay, the total number of nuclei from the selection cultures added to microscope slides, and the LI of cells from the non-selection culture
- for the cloning assay, the numbers of cells plated/well in non-selection and selection plates
- for both assays, the total number of MNCs assayed
- for both assays, the total number of evaluatable nuclei or cells in the assay
- for the cloning assay, the CEs (non-selection) as determined for each inoculum used for determining the overall non-selection CE for that assay.

6.6. Literature cited

1. R.J. Albertini, R.B. Hayes, Somatic cell mutations in cancer epidemiology, in: P. Toniolo, P. Boffetta, D.E.G. Shuker, N. Rothman, B. Hulka, N. Pearce (Eds.), IARC Scientific Publications 142 Application of biomarkers in cancer epidemiology, International Agency for Research on Cancer, Lyons, 1997, pp. 159–184, Human Mutation, Academic Press, New York, 1981, pp. 235–263.
2. J. Cole, T.R. Skopek, International Commission for Protection against Environmental Mutagens and Carcinogens, Working Paper 3, Somatic Mutant Frequency, Mutation Rates and Mutational Spectra in the Human Population in vivo, Mutat. Res. 304 (1994) 33–106.
3. B.A. Finette, T. Poseno, R.J. Albertini, V(D)J recombinase mediated *HPRT* mutations in peripheral blood of normal children, Cancer Res. 56 (1966) 1405–1412.
4. J.C. Fuscoe, L.J. Zimmerman, M.J. Lippert, J.A. Nicklas, J.P. O'Neill, R.J. Albertini, V(D)J recobminase-like activity mediates *HPRT* gene deletion in human fetal T-lymphocytes, Cancer Res. 51 (1991) 6001–6005.
5. D. Robinson, K. Goodall, R.J. Albertini, J.P. O'Neill, B.A. Finette, M. Sala-Trepat et al., An analysis of *HPRT* mutant frequency in circulating T-lymphocytes in the normal human population: a comparison of four databases, Mutat. Res. 313 (1994) 227–247.
6. A.D. Tates, T. Grummt, M. Tornqvist, P.B. Farmer, F.J. van Dam, H. van Mossel, H.M.

Shoemaker, S. Osterman-Golkar, C. Uebel, Y.S. Tang, A.H. Zwinderman, A.T. Natarajan, L. Ehrenberg, Biological and chemical monitoring of occupational exposure to ethylene-oxide, Mutat. Res. 250 (1991) 483–497.
7. C.B. Thomas, D.O. Nelson, P. Pleshanov, I. Vorobstova, L. Tureva, R. Jensen, I.M. Jones, Elevated frequencies of hypoxanthine-guanine phosphoribosyltransferase mutants are detected in Russian liquidators 6 to 10 years after exposure to radiation from the Chenobyl nuclear power plant accident, Mutat. Res. 439 (1999) 105–119.
8. D.J. Grdina, B.A. Carnes, D. Grahn, C.P. Sigdestad, Protection against late effectsof radiation by *S*-2-(3-aminopropylamino)-ethylphosphorothioic acid, Cancer Res. 51 (1991) 4125–4130.
9. D.J. Grdina, Y. Kataoka, I. Basic, J. Perrin, The radioprotector WR-2721 reduces neutron-induced mutations at the hypoxanthine-guanine phosphoribosyltransferase locus in mouse splenocytes when administered prior to or following irradiation, Carcinogenesis 13 (1992) 811–814.
10. A. Aidoo, L.E. Lyn-Cook, R.A. Mittelstaedt, R.H. Heflich, D.A. Casciano, Induction of 6-thioguanine-resistant lymphocytes in Fischer 344 rats following in vivo exposure to *N*-ethyl-*N*-nitrosourea and cyclophosphamide, Environ. Mol. Mutagen. 17 (1991) 141–151.
11. R.F. Branda, J.P. O'Neill, L.M. Sullivan, R.J. Albertini, Factors influencing mutation at the *HPRT* locus in T-lymphocytes: women treated for breast cancer, Cancer Res. 51 (1991) 6603–6607.
12. M.M. Ammenheuser, J.B. Ward, Jr., E.B. Whorton, Jr., J.M. Killian, M.S. Legator, Elevated frequencies of 6-thioguanine resistant lymphocytes in multiple sclerosis patients treated with cyclophosphamide: a prospective study, Mutat. Res. 204 (1988) 509–520.
13. S.M. Hou, F. deZwart, C.A. Warnock, C.R. Barnett, Y.A. Barnett, L. Celotti, M.Davies, W. Huettner, B. Lambert, A.D. Tates, Validation of the *HPRT* mutation assay using primary human T-cells — report from the EUCAHM ring tests, Mutat. Res. 431 (1999) 211–221.
14. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
15. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health, Biosafety in Microbiological and Biomedical Laboratories, HHS Publication No (CDC) 93-8395. 3rd edn. (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H. Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices, 2nd edn., ASM Press, Washington DC, 1995.
16. J.P. O'Neill, M.J. McGinniss, J.K. Berman, L.M. Sullivan, J.A. Nicklas, R.J. Albertini, Refinement of a T-lymphocyte cloning assay to quantify the in vivo thioguanine-resistant mutant frequency in humans, Mutagenesis 2 (1987) 87–94.
17. A.D. Tates, P.J. Boogaard, F. Darroudi, A.T. Natarajan, M.E. Caubo, N.J. van Sittert, Biological effect monitoring in industrial workers following incidental exposure to high concentrations of ethylene oxide, Mutat. Res. 329 (1995) 63–77.
18. R.J. Albertini, L.S. Sullivan, K.J. Berman, C.J. Greene, J.A. Stewart, J.M. Silveira, J.P. O'Neill, Mutagenicity monitoring in humans by autoradiographic assay for mutant T-lymphocytes, Mutat. Res. 204 (1988) 481–492.
19. M.M. Ammenheuser, D.A. Hastings, E.B. Whorton, Jr., J.B. Ward, Jr., Frequencies of *HPRT* mutant lymphocytes in smokers, nonsmokers, and former smokers, Environ. Molec. Mutagen. 30 (1997) 131–138.
20. P. Ostrosky-Wegman, R. Montero, C.C. deNava, R. Tice, R. Albertini, Use of BrdUrd labelling in the human lymphocyte HGPRT somatic mutation assay, Mutat. Res. Lett. 191 (1987) 211–214.
21. R. Montero, H. Norppa, K. Autio, C. Lindholm, P. Ostrosky-Wegman, M. Sorsa, Determination of 6-thioguanine resistant lymphocytes in human blood by immunohistochemical antibromodeoxyuridine staining, Mutagenesis 6 (1991) 169–170.
22. G. Hageman, I. Welle, R. Stierum, H. Albering, J. Kleinjans, Detection of thioguanine resistant human peripheral blood lymphocytes using 5-

bromodeoxyuridine labelling in combination with immunocytochemical staining, Mutagenesis 8 (1993) 495–501.
23. F. Vigagni, P. Hrelia, G. Cantelli-Forti, M. Sorsa, H. Norppa, An improved assay to detect 6-thioguanine resistant mutants in human lymphocytes, J. Experimen. Clinic. Cancer Res. 14 (1995) 171–173.
24. J. Cole, M.H.L. Green, S.E. James, L. Henderson, H. Cole, A further assessment of factors influencing measurements of thioguanine-resistant mutant frequency in circulating T-lymphocytes, Mutat. Res. 204 (1988) 493–507.
25. J.P. O'Neill, J.A. Nicklas, T.C. Hunter, O.B. Batson, M. Allegretta, M.T. Falta et al., The effect of T-lymphocyte ''clonality'' on the calculated *HPRT* mutation frequency occurring in vivo in humans, Mutat. Res. 313 (1994) 215–225.
26. R.J. Albertini, E.F. Allen, A.S. Quinn, M.R. Albertini, Human somatic cell mutation: in vivo variant lymphocyte frequencies as determined by 6-thioguanine resistance, in: E.B. Hook, I.H. Porter (Eds.), Population and Biological Aspects of Human Mutation, Academic Press, New York, 1981, pp. 235–263.

7. Chromosome aberrations in lymphocytes

7.1. Introduction

7.1.1. Purpose of the guideline

To provide concise guidance for various methodologies used for evaluating structural and numerical chromosomal aberrations in peripheral blood lymphocytes sampled from individuals potentially exposed to clastogens or aneuploidogens. This guideline must be read in conjunction with the General Introduction (Section 2), which contains information on such issues as study design, ethics, and appropriate statistical considerations.

7.1.2. Principles of the assay

Structural chromosomal aberrations result from: (A) direct DNA breakage; (B) replication on a damaged DNA template; (C) inhibition of DNA synthesis [1–4] and other mechanisms (e.g., topoisomerase II inhibitors [5–7]). Very few agents (e.g., ionising radiation, bleomycin) induce direct DNA breakage. These agents induce, at the time of exposure, chromosome-type chromosomal aberrations (involving both chromatids of a chromosome) in cells in the G_0/G_1 phase of the cell cycle and chromatid-type (involving only one chromatid of a chromosome) chromosomal aberrations in cells in the S/G_2 phase [4,8,9]. Operationally, these agents are classified as S-phase-independent clastogens. In vivo exposure of blood G_0 lymphocytes to these clastogens results in chromosome-type chromosomal aberrations that persist for the lifetime of the cell and can be detected during the first mitotic division in vitro. The second mechanism is the one most likely to be applicable to many human population monitoring studies as it represents the most common process by which clastogenic chemicals induce chromosomal aberrations [10]. Operationally, these agents are classified as S-phase-dependent clastogens because the damaged cells must enter S-phase before chromosomal damage (chromatid-type chromosome aberrations) is elicited. For these agents, aberrations induced in blood lymphocytes arise predominantly in vitro as chromatid-type during the first post-mitogenic stimulation S-phase in response to the DNA damage induced in vivo. S-phase-dependent chromosomal aberrations can be induced in vivo in proliferating lymphocytes, but the frequency of such events is expected to be low due to the (normally) infrequent occurrence of cycling lymphocytes, and selection against cells with non-stable chromosomal aberrations. As the induction of chromosomal damage by the third mechanism, i.e., via interference with DNA synthesis (but not via direct DNA damage) occurs only when the appropriate agents (e.g., hydroxyurea) are present during S-phase in vivo, the third mechanism is unlikely to be responsible for chromosomal aberrations detected in lymphocytes in vitro. However, in exposure situations where relatively high concentrations of an agent of this type are present in blood and whole blood cultures are used, an increased frequency of chromosomal aberrations in vitro due to this mechanism is theoretically possible.

Numerical chromosomal aberrations (i.e., aneuploidy, polyploidy) refer to changes in chromosome number that occur due to abnormal cell division.

Cells are classified as aneuploid when they contain up to a few more (hyperploid) or a few less (hypoploid) chromosomes than the normal complement (46 for humans); cells are classified as polyploid when they contain multiples of the normal complement i.e., 4 N, 8 N, etc. The mechanisms by which aneuploidy can occur include damage to the mitotic spindle and associated elements, damage to chromosomal sub-structures, alterations in cellular physiology, and mechanical disruption [11]. The mechanisms by which polyploidy occurs, are less clear [12].

Structural and numerical chromosomal aberrations are most commonly scored in proliferating cells arrested at metaphase using a tubulin polymerisation inhibitor (e.g., Colcemid®, colchicine). For chronic exposure studies, the frequency of structural chromosomal aberrations detected in mitogen-stimulated peripheral blood lymphocytes sampled from individuals exposed to clastogenic agents depends on the level, duration, and frequency of exposure and the clastogenic mechanism involved. For S-dependent agents, the frequency is further modulated by the in vivo balance between the induction of DNA damage and its repair, and by the extent of repair occurring in vitro between sampling and the first S-phase and the efficiency with which the induced lesions are expressed as chromosomal damage. For acute exposure studies and for chronic exposure studies where the exposure has ceased, the extent to which the DNA lesions persist in lymphocytes between exposure and blood collection is a critical determinant in the level of chromosomal damage detected in vitro. Following an in vivo exposure to aneugenic (aneuploidogenic) agents, aneuploid or polyploid lymphocytes are unlikely to be induced in vitro unless damage to the kinetochore and/or centromeric region is the mechanism involved (unless the chemical persists to the in vitro situation). If induced in vitro, such cells must have divided at least once during culturing. However, aneuploid or polyploid cells arising in vivo can be ascertained analysing first generation metaphase cells.

In addition to classical cytogenetic methods for scoring chromosomal aberrations in metaphase cells, structural and numerical damage may be detected in metaphase- and interphase-stage lymphocytes (or interphase-stage epithelial cells) using fluorescence in situ hybridisation (FISH) [13–22]. This technique provides increased efficiency and specificity for identifying certain kinds of chromosomal aberrations (i.e., chromosome-type structural rearrangements, stable symmetrical rearrangements derived from chromatid-type chromosome aberrations, hyperploidy) induced in vivo.

Because of the generally ready availability of blood and the need for cycling cells, the focus of this guideline is on blood lymphocytes.

7.1.3. Definitions

An aneugen (aneuploidogen) is an agent capable of inducing aneuploidy. Chromatid-type aberrations are aberrations that involve one sister chromatid of any one chromosome or more chromosomes. Chromosome-type aberrations involve the same locus on both sister chromatids on one or multiple chromosomes. A clastogen is an agent that induces chromosomal aberrations. S-independent agents are agents that induce DNA strand breaks which are expressed in the first metaphase post-exposure as chromosome-type chromosomal aberrations for cells damaged in G_1 and chromatid-type chromosomal aberrations for cells damaged in S/G_2. S-dependent agents are agents which induce DNA damage which require DNA synthesis before being expressed in the first metaphase post-exposure (or if the insult took place in G2, in the second metaphase) as chromatid-type chromosomal aberrations.

7.1.4. Significance of the endpoint and application in risk assessment

There is experimental and epidemiological evidence for the involvement of structural and numerical chromosomal aberrations in carcinogenesis [23–25]. Also, at the group level, an increased frequency of structural chromosomal aberrations in peripheral blood lymphocytes is associated with an increased overall risk for cancer [26–28]. No information is available on the possible association of an elevated frequency of aneuploidy in lymphocytes with an increased cancer risk.

7.2. Population characteristics

Important characteristics to consider for the exposed and referent populations are discussed in Sec-

tion 2. As presented in this section, specific information on socio-economic status and lifestyle (e.g., smoking, recreational/medical drug usage), health status, and endogenous factors, such as age and sex, need to be obtained for each subject. When possible, individuals with genetic polymorphisms associated with increased susceptibility to genotoxic agents can be identified. Data demonstrating that these factors modulate the frequency of structural and numerical chromosomal aberrations in blood lymphocytes are still scanty, but suggest that, for instance, polymorphisms of glutathione *S*-transferases M1 and T1 may affect chromosome aberration frequencies [29–34]. A concurrent referent population is required, except in the case of acute exposure to ionising radiation, for which historical control data are available [35].

As discussed in Section 2, the number of subjects needed for a particular study design depends, for a desired power, on the magnitude of the expected effect and the known variability of the measured endpoint [36]. Information on optimal population size characteristics for structural chromosomal aberrations in human blood lymphocytes should always be considered during study design. Appropriate software programs are available to assist the investigator in making these calculations (see Section 2). Information on population size calculations for structural chromosomal aberrations can be also be found [37–39]. Similar information on the number of subjects needed for a study designed specifically to assess for aneuploidy has not been published.

Where the number of subjects to evaluate cannot be determined prior to study implementation or in situations where the appropriate number of subjects are not available, the precision in the risk estimate should be considered in evaluating the results.

7.3. Sample timing

The optimal sampling time for peripheral blood lymphocytes is during a long-term chronic exposure when the induction and repair of DNA damage in lymphocytes (or other cell population studied) is presumed to be at steady-state. In most cases, this sample time maximises the likelihood of identifying an exposure to a clastogenic agent. For the collection of blood after an acute exposure or after termination of a chronic exposure to an S-dependent clastogenic agent, the optimal collection time for detecting induced damage capable of being expressed as chromosomal damage in vitro is within a few hours of exposure termination. The frequency of DNA damage in lymphocytes decreases with increasing time between exposure and sampling due to DNA repair processes, the loss of heavily damaged cells from the blood through apoptosis or necrosis, and cell turnover. In general, blood samples should be collected within 2 days after an acute exposure; they may be collected at a later time but with an increasing likelihood that a positive clastogenic effect will not be detectable.

There is little information on the detection of aneuploidy in humans after an in vivo exposure, and the applicable methods are still in a developmental stage. Therefore, this guideline provides more or less general information on aneuploidy assays. Detailed recommendations on, e.g., cell types to be studied, techniques to be used, and the choice between direct sampling and cell culture cannot yet be given. In chronic exposure, cultured or uncultured lymphocytes may be analysed. The assessment of aneuploidy induced in vivo during an acute exposure may be possible if FISH on short-lived cell populations is utilised (epithelial cells, granulocytes).

7.4. Methods

Descriptions of this assay and its variants can be found in Refs. [38,39].

7.4.1. Quality assurance

A system for recording the collection, transport, and storage of samples should be established to ensure that sample integrity is maintained.

The use of coded samples where scoring is conducted without knowledge of exposure information is critical for eliminating potential bias.

Although seldom employed in past human population studies, it is important to include a quality assurance step to demonstrate the adequacy of the methodology used in the study. In some situations, subsets of the population (e.g., smokers versus nonsmokers) may serve, retrospectively, to demonstrate the adequacy of the methodology. However, this approach is unlikely to be generally useful. It would be more appropriate to include, as part of the study design, a quality assurance control among the sam-

ples being analysed. This can be accomplished, for example, by including fixed, stored cells or previously prepared slides with a known induced low frequency of chromosomal aberrations in the study design. This approach assesses the ability of the scorers to appropriately detect chromosomal damage and must be accomplished in such a manner that the identity of the samples/slides remains unknown to the scorer.

In addition to these internal quality assurance steps, inclusion of external quality assurance steps should be considered in the study design. External steps could involve the sharing of replicate blood samples with an independent laboratory or the independent re-scoring by another laboratory of a certain fraction of the total slides scored.

7.4.2. Blood collection

Because of the risk for infectious diseases (hepatitis, HIV), appropriate precautions must always be followed when handling blood samples (see Ref. [40,41]).

At the time of collection, blood is mixed with heparin to prevent coagulation. It is recommended that blood samples be processed within 24 h of collection but this may not always be possible, and lymphocyte from blood stored for at least several days can be cultured successfully. However, for S-dependent clastogenic agents, increasing the time between collection and processing is likely to be accompanied by a decline in DNA damage due to DNA repair. To minimise the loss of damage due to DNA repair processes, it would be useful to refrigerate the blood samples at below 8°C between collection and processing. In many situations, blood collection will occur at sites distant from the cytogenetics laboratory and transportation will be necessary. If blood samples are shipped commercially, a minimum–maximum thermometer in the shipping container will provide a measure of the temperature range experienced. If air transportation is used, the blood should not be subjected to X-irradiation at security checkpoints. A sheet of X-ray film can be included in the shipping package to verify if such an exposure occurred, although the doses involved in this type of X-irradiation are low and probably not detectable by chromosome aberration analysis. If samples are shipped internationally, appropriate permits should be obtained in advance and included in the shipment to void delays at customs. Most importantly, blood from both exposed and non-exposed individuals must be handled in the same manner. Where blood collection is protracted across time, a balanced collection design must be used.

7.4.3. Description

7.4.3.1. Culturing. It may be useful to determine the total white cell count and to prepare a slide for a differential count from each blood sample just prior to establishing the cultures. This is required for cultures of purified lymphocytes to standardise the number of lymphocytes in the cultures, but may also be useful during data interpretation; e.g., induced leucopenia indicates loss of white cells due to exposure. Selection of culture medium, serum type and concentration and the use of whole blood versus purified lymphocyte cultures are arbitrary and should be based on experience. Culture medium (e.g., Dulbeccos's minimal essential medium without added nonessential amino acids) or other culture conditions associated with the expression of fragile sites should be avoided. It is recommended that the same lot of media, serum, and other reagents be used, whenever possible, throughout a study.

Whole blood or isolated mononuclear cells are added to culture medium containing phytohaemagglutinin (PHA) (if justified, other mitogens maybe used), and incubated at 37 ± 1°C. If the culture containers are not sealed tightly, the cultures should be maintained in humidified incubator equilibrated with 5% CO_2. At least duplicate cultures are established for each blood sample, but more cultures may be established if deemed useful.

7.4.3.2. Termination. The single most important aspect of culturing blood samples for the analysis of chromosomal aberrations in human biomonitoring studies is the harvest time [38]. The maximal frequency of chromosomal aberrations in lymphocytes collected from exposed individuals occurs in first-generation, post-exposure metaphase cells. The frequency of chromosomal aberrations (except balanced translocations) is decreased greatly in second-generation metaphase cells due to the dilution of damage among daughter cells and the loss of damaged cells

from the proliferating cell population. Generally, in an attempt to limit scoring to first-generation metaphase cells, most laboratories harvest blood cultures at 48-h post PHA-stimulation. This harvest time, or a shorter harvest time, is acceptable if it can be ascertained that, in most donors, the great majority of metaphase cells are in their first division. However, the proportion of first- to second-generation metaphase cells at this time can vary greatly among individuals and depends on subject-to-subject variability in mitogen-dependent growth characteristics under the culture conditions used [42]. The method most commonly used to ensure that first-generation metaphase cells only are scored for chromosomal damage is based on adding bromodeoxyuridine (BrdU) to the culture medium prior to the onset of the first round of DNA replication in vitro [43]. Incorporation of BrdU into the DNA of replicating cells allows for the unequivocal identification of first versus subsequent generation metaphase cells. The concentration of BrdU used should be one demonstrated not to affect cell kinetics or chromosome aberration frequency. Where alterations in growth characteristics might be of concern and where resources permit, multiple sample times (e.g., 48–60 h) can be used to ensure an adequate number of first generation metaphase cells for scoring.

Cell cultures are treated with colchicine or Colcemid® generally one to two h prior to harvesting. The cells are harvested and processed through treatments with a hypotonic solution (e.g., 0.075 mol/l KCl) and fixative (e.g., 3:1 methanol:glacial acetic acid). Once fixed, cells can be stored for several years in the fixative solution at 4°C to −20°C.

7.4.3.3. Slide preparation. Using standard methods, slides are prepared. For the scoring of chromosomal aberrations in first-generation metaphase cells, cells are stained with a method (e.g., fluorescence plus Giemsa) that provides for the differential staining of metaphase cells based on the pattern of incorporated BrdU. If BrdU is not used, other appropriate stains (e.g., Giemsa) are applied. After staining, the slides are coverslipped to protect the cells and then stored for scoring. For an analysis based on on FISH, unstained slides are stored at −20°C until processed using standard methodology [44].

7.4.3.4. Scoring. Uniform criteria for scoring chromosomal aberrations, such as that published by Savage [45] or contained in the International System for Human Cytogenetic Nomenclature [46], should be used. Chromosomal aberrations are classified based on the number of sister chromatids and breakage events involved (i.e., chromosome-type versus chromatid-type; simple versus complex). Simple chromosomal aberrations (e.g., breaks, deletions) are hypothesised to involve only a single breakage event while complex aberrations (rearrangements within and between chromosomes) involve multiple breakage and misrepair events. Information on the types of chromosomal aberrations detected is important for interpretation of the results. Metaphase analysis should be conducted by a trained and experienced observer completely familiar with the chromosome aberration analysis. Where multiple scorers are involved, a balanced scoring design must be used.

Slides must be scored without knowledge of the exposure group; slides from the quality assurance standard should be included within this set of slides. Where possible, the number of metaphase cells to be analysed per subject should be based upon the frequency of cells with damage in the referent population, or on expectations based on historical control data. In general, the goal is to obtain as few subjects as possible with zero class data while avoiding unnecessary costs. Based on an average population rate of at least 1% for cells with chromosomal damage, a minimum of 200 metaphase cells per subject should be scored [37]. For more information, see Section 2.

Metaphase cells for analysis should be identified under low power magnification and intact cells with well-spread chromosomes should be examined using a high power objective (at least 63 × magnification). Only metaphases containing 45–47 centromeres are analysed. Gaps (achromatic regions) are recorded separately. Criteria for distinguishing gaps from breaks and simple from complex aberrations have been published [45–48].

In addition to classical aberration scoring methods, techniques based on FISH can be used to score for structural aberrations. Cells are stained with whole chromosome paint probes for from one to six pairs of chromosomes using different coloured probes, and the number of metaphase cells equal to 200–1000 (depending on the average population rate of the

aberrations scored) of the whole genomic equivalents is scored for clastogenic damage [21,22,48]. When BrdU is used but first division cells have not been identified at or before the analysis, the slides can be restained after scoring and the same metaphase cells evaluated for replicative history (i.e., by FPG staining or BrdU antibody staining) to identify any second-generation metaphase cells [17]. These cells are eliminated from the database, and if the proportion of second-generation metaphase cells is appreciable (e.g., $> 20\%$), more metaphase cells are scored using FISH in an attempt to obtain the required number of genomic equivalents.

Aneuploid cells can be detected in metaphase preparations prepared for structural chromosome aberration analysis. However, because of concern about the loss of chromosomes from metaphase cells associated with slide preparation, only hyperploidy is used as a measure of aneuploidy. Furthermore, as the assessment of aneuploidy is often considered a minor component of structural chromosomal aberration studies, only limited attention has been given to developing the most appropriate study design for this endpoint. The appropriate number of metaphase cells to be scored per subject should be based on the average rate of hyperploidy in the general referent population. An assessment of polyploidy should use the same approach.

The evaluation of aneuploidy in classical metaphase preparations is time consuming and labour intensive. The FISH technique using whole chromosome specific DNA libraries allows for a rapid assessment of the number of labelled chromosomes in each metaphase, enabling the efficient scoring of a large number of cells [18,49]. Several chromosome pairs can simultaneously be evaluated using multicoloured FISH. Again, only hyperploid events are scored due to the potential for technical artefacts resulting in the loss of a hybridisation signal.

Efficient methods for the analysis of aneuploidy in interphase cells have also been developed, using centromeric FISH techniques to identify specific chromosomes in nuclei and whole chromosomes in MN (see Section 8). Most of the studies conducted to date utilising chromosome-specific centromeric DNA probes have used mitogen-stimulated lymphocytes although unstimulated lymphocytes or other cells (e.g., buccal mucosa) can be used [14,15,19,20,49–52]. To avoid scoring polyploids as hyperploids, it may be necessary to use centromeric DNA probes of two different chromosomes labelled with different fluorescent dyes [14,52]. At present, no recommendation can be given regarding the chromosomes preferred to be included in such analysis. It appears that the X chromosome has a high propensity to malsegregate in male and female lymphocytes [52–54], but it is not known if also induced aneuploidy could preferentially affect this chromosome. If this were the case, the inclusion of the X chromosome would be recommendable. If, on the other hand, induced aneuploidy occurred at a similar rate for the X chromosome and autosomes, the X chromosome would be a bad choice, due to its high background frequency of aneuploidy. The appropriate number of metaphase or interphase cells to be scored per subject should be based on the average rate of hyperploidy, as assessed using this technique, in the general referent population, and the number of chromosome probes used.

Simultaneous detection of clastogenic and aneugenic events in interphase or metaphase cells can be achieved by using tandem labelled probes specific for two adjacent regions of a chromosome [16,55]. The probes are labelled with two different fluorochromes (e.g., red and green). Aneuploidy-related events will lead to additional signals of both probes; clastogenic events will lead to breakage between and within the labelled regions. Again, the number of cells scored should be based on the appropriate number of genomic or numerical equivalents. The difficulty inherent in this method, at least for S-dependent clastogenic events completed in vitro, is in ensuring that the interphase cells scored have completed at least one S-phase. Several methods, including the use of BrdU to identify post S-phase cells or scoring binucleate cells induced by cytochalasin B can be used to limit scoring to the population of cells at risk for chromosomal damage. Chromatid fragments are not separated from the intact sister chromatid before anaphase and may not be detected in G2-phase cells. The system should work in binucleate cells, where one can see where the fragment migrated — it can also be in a micronucleus. Another yet unclear issue with this method is whether chromatin organisation in interphase nuclei (S-phase) could influence the results.

In addition to measures of clastogenic and aneugenic damage, information obtained on the relative frequencies of first-generation, second-generation, and subsequent generation metaphase cells, based on BrdU-dependent differential staining patterns, can be used to evaluate cell proliferation kinetics in cultured lymphocytes [38,56,57]. Such information, although not indicative of genotoxic damage, is useful in identifying exposure-induced alterations in mitogen responsiveness and/or the subsequent rate of cell division [17,56,58].

7.4.3.5. Data recording. For each cell containing structural or numerical chromosomal damage, a detailed description and number of the different types of structural aberrations detected and the microscope co-ordinates should be recorded. When a metaphase finder is available for the analysis, the location (and in some models image) of all cells scored and the findings can be stored electronically, facilitating optimal quality assurance. When non-permanent slides (fluorescence techniques) are used, it is a good practice to store the images of abnormal cells either as photographs or electronically (in the memory of an imaging station).

7.4.4. Sources of assay variation

The type of culture medium used can affect cell cycle progression, and, some types of media can increase the expression of fragile chromosome sites.

The type of mitogen influences the proliferation rate of lymphocytes.

The stability of the temperature control of incubators, and the temperature variation within the incubators can affect differentially the growth of cultures.

The type/batch of serum may influence the frequency of chromosomal aberrations.

A lack of uniform collection, shipping, processing, and storage of samples will increase technical variability.

Variation among scorers is important in all microscopy, including CA analyses. Variation can be decreased by having only one scorer, and, when several scorers are used, reduced by ensuring a balanced scoring design. Thus, e.g., each scorer should analyse the same number of cells from the slides of all subjects (the same cells not being scored twice) rather than one scorer analysing all cells from some subjects and another scorer all cells from other subjects. Variation can also be reduced by ensuring that the scoring is not unnecessarily protracted across time. Each scorer should be experienced so that scoring criteria do not evolve during the analysis.

7.4.5. Statistical analysis

As described in Section 2, data should be evaluated according to the statistical plan of the study. For each subject, the total aberration frequency and the percentage of damaged cells are calculated, in the classical method including and excluding gaps. For aneuploidy, the percentage of hyperploid and polyploid cells are considered separately. There is no unified agreement on the proper statistical methods to be used. However, the unit of exposure is the individual and not the cell; pooling cells across individuals within a study group eliminates a critically important source of variability. The choice of statistical method to be used may depend on the distribution of the data. If the distribution of the data is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, parametric methods based upon normal distribution can also be used, when appropriate, such as after relevant data transformation. When needed, zero values can be transformed [59]. Statistical advice may be needed to determine whether parametric or non-parametric methods are more appropriate for a particular study. The relative and absolute effects of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion by estimating the precision of the effects seen in the study. Stratified analyses can be used for exploring the effect modification by factors such as age, sex, smoking, other exposures, etc. More sophisticated multivariate methods may be necessary for investigating the effect of factors that either may modify or be confounded with the effects [60,61]. Wherever possible, the analysis should be based on individual exposure information. Consideration for the effect of multiple comparisons on the false positive rate should be included in the statistical analysis.

The quality assurance standard is compared against the appropriate concurrent control cultures using appropriate statistics.

7.4.6. Interpretation of results

Criteria for identifying a positive response (i.e., increased levels of chromosomal aberrations in an exposed group compared with the referent population; a dose response relationship if variable exposure levels have been established) must be identified prior to implementing the study.

A statistically significantly higher frequency of chromosomal aberrations among the exposed population in comparison with the referent population is considered a positive result (i.e., the agent of interest induced clastogenic or aneugenic damage in lymphocytes of exposed individuals). A positive result is supported by a dose–response relationship when individuals with different exposure levels are considered. Presentation of the results in terms of the size of the effect with confidence intervals may also help with the interpretation. Maximum reliability that the exposure results in an increased frequency of chromosomal damage requires reproducible results in independent studies. It is appreciated that independent studies are not always possible (e.g., in the case of accidents).

Information on the types of aberrations induced (S-independent versus S-dependent) can be used to identify to some extent the nature of the clastogenic damage induced (i.e., strand breaks versus base damage, respectively).

The lack of a statistically significant increase in chromosomal damage in lymphocytes indicates that under the exposure conditions evaluated and for the calculated power of the study, the exposure did not result in a significant increase in chromosomal damage in the cell population evaluated.

The results of the quality assurance control are used to demonstrate the adequacy of the methodology and the scorers in identifying chromosomal damage.

7.5. Report

All data should be presented in tabular form and include all observed results (see below). This will allow comparisons between laboratories, re-analyses of results by outside reviewers, and provide a database for baseline background values and for exposures to physical and chemical agents. All results of the study should be reported with none omitted, and the definition and handling of ''outliers'' described.

The study report should include the following minimal information:

- copy of the questionnaire used to determine personal characteristics
- statement of ethical approval
- characteristics of the exposed and referent populations
- exposure information and analytical methods used
- blood collection procedures, including timing in relationship to exposure, transportation, storage, and chain-of-custody
- method for coding samples prior to analysis
- culture conditions (if used), including medium composition, mitogen type and concentration, CO_2 concentration, incubation duration, colchicine or Colcemid® concentration and duration
- number of cells scored per culture or sample and number of replicates
- number of metaphases or cells analysed (data given separately for each subject)
- criteria for scoring aberrations
- criteria for cell kinetic analysis (if conducted)
- number of scorers, the scoring design, and nomenclature system used
- type and number of aberrations for each subject
- kinetic data for each subject (if scored)
- tabulated results for each subject
- statistical methods
- discussion and interpretation of results
- internal and external quality assurance procedures.

7.6. Literature cited

1. M.A. Bender, Relationship of DNA lesions and their repair to chromosomal aberration production, in: W.M. Generoso, M.D. Shelby, F.J. de Serres (Eds.), DNA Repair and Mutagenesis in Eukaryotes, Plenum, New York, 1980, pp. 245–265.
2. M.A. Bender, H.G. Griggs, J. Bedford, Mecha-

nisms of chromosomal aberration production. III. Chemicals and ionising radiation, Mutat. Res. 23 (1974) 197–212.
3. H.J. Evans, D. Scott, The induction of chromosome aberrations by nitrogen mustard and its dependence on DNA synthesis, Proc. R. Soc. London Ser. B Biol. Sci. 173 (1969) 491–512.
4. B.A. Kihlman, Caffeine and Chromosomes, Elsevier, Amsterdam, 1977, 68 pp.
5. K. Wassermann, Intragenomic heterogeneity of DNA damage formation and repair: a review of cellular responses to covalent drug DNA interaction, Crit. Rev. Toxicol. 24 (1994) 281–322.
6. L.R. Ferguson, B.C. Baguley, Mutagenicity of anticancer drugs that inhibit topoisomerase enzymes, Mutat. Res. 355 (1996) 91–101.
7. L.R. Ferguson, Inhibitors of topoisomerase II enzymes: a unique group of environmental mutagens and carcinogens, Mutat. Res. 400 (1998) 271–278.
8. K.E. Buckton, H.J. Evans (Eds.), Methods for the analysis of human chromosome abberations, WHO Geneva, World Health Organization, Geneva, 1973, 66 pp.
9. H.J. Evans, M.L. O'Riordan, Human peripheral blood lymphocytes for the analysis of chromosome aberrations in mutagen test, Mutat. Res. 31 (1975) 135–148.
10. A.T. Natarajan, G. Obe, Screening of human populations for mutations induced by environmental pollutants: use of human lymphocyte system, Ecotoxicol. Environ. Safety, 4 (1980) 468–481.
11. M. Oshimura, J.C. Barrett, Chemically induced aneuploidy in mammalian cells: mechanisms and biological significance in cancer, Environ. Mutagen. 8 (1986) 129–159.
12. I. de G. Mitchell, T.R. Lambert, M. Burden, J. Sunderland, R.I. Porter, J.B. Carlton, Is polyploidy an important genotoxic lesion? Mutagenesis 10 (1995) 79–83.
13. D. Pinkel, T. Straume, J.W. Gray, Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridisation, Proc. Natl. Acad. Sci. U. S. A., 83 (1986) 2934–2938.
14. D.A. Eastmond, D. Pinkel, Detection of aneuploidy and aneuploidy-inducing agents in human lymphocytes using fluorescence in situ hybridisation with chromosome-specific DNA probes, Mutat. Res. 234 (1990) 303–318.
15. D.A. Eastmond, M. Schuler, D.S. Rupa, Advantages and limitations of using fluorescence in situ hybridisation for the detection of aneuploidy in interphase human cells, Mutat. Res. 348 (1995) 153–162.
16. D.S. Rupa, L. Hasegawa, D.A. Eastmond, Detection of chromosomal breakage in the 1cen–1q12 region of interphase human lymphocytes using multicolor fluorescence in situ hybridisation with tandem DNA probes, Cancer Res. 55 (1995) 640–645.
17. J.J.W.A. Boei, S. Vermeulen, A.T. Natarajan, Detection of chromosomal aberrations by fluorescence in situ hybridisation in the first three postirradiation divisions of human lymphocytes, Mutat. Res. 349 (1996) 127–135.
18. M.T. Smith, L. Zhang, Y. Wang, R.B. Hayes, G. Li, J. Wiemels et al., Increased translocations and aneusomy in chromosomes 8 and 21 among workers exposed to benzene, Cancer Res. 58 (1998) 2176–2181.
19. J. Surrallés, K. Autio, L. Nylund, H. Järventaus, H. Norppa, T. Veidebaum et al., Molecular cytogenetic analysis of buccal cells and lymphocytes from benzene exposed workers, Carcinogenesis 18 (1997) 817–823.
20. S.A. Salama, M. Serrana, W.A. Au, Biomonitoring using accessible human cells for exposure and health assessment, Mutat. Res. 436 (1999) 99–112.
21. K. Matsumoto, M.J. Ramsey, D.O. Nelson, J.D. Tucker, Persistence of radiation-induced translocations in human peripheral blood determined by chromosome painting, Radiat. Res. 149 (1998) 602–613.
22. M.J. Ramsey, D.H. Moore, II, J.F. Briner, D.A. Lee, L. Olsen, J.R. Senft, J.D. Tucker, The effects of age and lifestyle factors on the accumulation of cytogenetic damage as measured by chromosome painting, Mutat. Res. 338 (1995) 95–106.
23. F. Mitelman, Chromosomes, genes, and cancer [editorial], CA Cancer J. Clin. 44 (1994) 133–135.
24. D.A. Rew, Significance of aneuploidy, Br. J. Surg. 81 (1994) 1416–1422.

25. F. Mitelman, F. Mertens, F. Johansson, A break point map of recurrent chromosomal rearrangements in human neoplasia, Nature (Genetics), 15 (1997) 417–474.
26. L. Hagmar, A. Brogger, I.-L. Hansteen, S. Heim, B. Högstedt, L. Knudsen et al., Cancer risk in humans predicted by increased levels of chromosomal aberrations in lymphocytes: Nordic Study Group on the Health Risk of Chromosome Damage, Cancer Res. 54 (1994) 2919–2922.
27. S. Bonassi, A. Abbondandolo, L. Camurri, L. Dal Prá, M. De Ferrari, F. Degrassi et al., Are chromosome aberrations in circulating lymphocytes predictive of a future cancer onset in humans? Preliminary results of an Italian cohort study, Cancer Genet. Cytogenet. 79 (1995) 133–135.
28. L. Hagmar, S. Bonassi, U. Stromberg, A. Brogger, L.E. Knudsen, H. Norppa, C. Reuterwall, Chromosomal aberrations in lymphocytes predict human cancer: a report from the European Study Group on Cytogenetic Biomarkers and Health (ESCH), Cancer Res. 58 (1998) 4117–4121.
29. M. Sorsa, S. Osterman-Golkar, K. Peltonen, S.T. Saarikoski, R. Sram, Assessment of exposure to butadiene in the process industry, Toxicology 113 (1996) 77–83.
30. R. Scarpato, A. Hirvonen, L. Migliore, G. Falck, H. Norppa, Influence of GSTM1 and GSTT1 polymorphisms on the frequency of chromosome aberrations of smokers andpesticide-exposed greenhouse workers, Mutat. Res. 389 (1997) 227–235.
31. N. Conforti-Froes, R. el-Zein, S.Z. Abdel-Rahman, J.B. Zwischenberger, W.W. Au, Predisposing genes and increased chromosome aberrations in lung cancer cigarette smokers, Mutat. Res. 379 (1997) 53–59.
32. S. Landi, H. Norppa, G. Frenzilli, G. Cipollini, I. Ponzanelli, R. Barale, A. Hirvonen, Individual sensitivity to cytogenetic effects of 1,2:3,4-diepoxybutane in cultured human lymphocytes: influence of glutathione *S*-transferase M1, P1 and T1 genotypes, Pharmacogenetics 8 (1998) 461–471.
33. R. Srám, P. Rossner, K. Peltonen, K. Podrazilova, G. Mrackova, N.A. Demopoulos et al., Chromosomal aberrations, sister-chromatid exchanges, cells with high frequency of SCE, micronuclei and comet assay parameters in 1, 3-butadiene-exposed workers, Mutat. Res. 419 (1998) 145–154.
34. L.E. Knudsen, H. Norppa, M.O. Gamborg, P.S. Nielsen, H. Okkels, H. Soll-Johanning et al., Chromosomal aberrations induced by urban air pollution in humans: influence of DNA repair and polymorphisms of glutathione *S*-transferase M1 and *N*-acetyltransferase 2, Cancer Epidemiol. Biomarkers Prev. 8 (1999) 303–310.
35. IAEA, Biological Dosimetry: Chromosome aberration analysis for dose assessment, International Atomic Energy Agency, Vienna, 1986, IAEA Technical Report 260, 68 pp.
36. J.D. Tucker, D.H. Moore, II, The importance of age and smoking in evaluating adverse cytogenetic effects of exposure to environmental agents, Environ. Health Perspect. 104 (Suppl. 3) (1996) 489–492.
37. E.B. Whorton, Jr., D.E. Bee, D.J. Kilian, Variations in the proportion of abnormal cells and required sample sizes for human cytogenetic studies, Mutat. Res. 64 (1979) 79–86.
38. A.V. Carrano, A.T. Natarajan, Considerations for population monitoring using cytogenetic techniques, International Commis sion for protection against environmental mutagens and carcinogens, ICPEMC Publication No. 14, Mutat. Res. 204 (1988) 379–406.
39. International Programme on Chemical Safety (IPCS), Guidelines for the study of genetic effects in human populations, World Health Organization, Geneva, Switzerland, 1985, Environmental Health Criteria 46, 126 pp.
40. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
41. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health, Biosafety in Microbiological and Biomedical Laboratories. HHS Publication No (CDC) 93-8395. 3rd edn. (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H.

Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices, 2nd edn., ASM Press, Washington DC, 1995.
42. G. Obe, B. Beek, The human leukocyte test system, in: F. de Serres, A. Hollaender (Eds.), Chemical Mutagenesis, 1982, pp. 337–400.
43. P. De Boer, P.P.W. van Buul, R. van Beek, F.A. van der Hoeven, A.T. Natarajan, Chromosomal radiosensitivity and karyotype in mice using cultured peripheral blood lymphocytes and comparison with this system in man, Mutat. Res. 42 (1977) 379–394.
44. A.T. Natarajan, S.J. Santos, F. Darroudi, V. Hagjidikova, S. Vermeulen, M. Van den Berg et al., 137 Cesium-induced chromosome aberrations analysed by fluorescence in situ hybridisation: eight years follow up of the Goiania radiation accident victims, Mutat. Res. 400 (1998) 299–312.
45. J.R.K. Savage, Classification and relationships of induced chromosomal structural changes, J. Med. Gen. 12 (1975) 103–122.
46. D.G. Harnden (Ed.), ISCN, An international system for Human Cytogenetic Nomenclature, S. Karger, Switzerland, 1985.
47. A. Brøgger, R. Norum, I.-L. Hansteen, K.O. Clausen, K. Skårdal, F. Mitelman et al., Comparison between five Nordic laboratories on scoring of human lymphocyte chromosome aberrations, Hereditas 100 (1984) 209–218.
48. J.R.K. Savage, J.D. Tucker, The nomenclature system for FISH-painted chromosome aberrations, Mutat. Res. 366 (1996) 153–161.
49. L. Zhang, N. Rothman, Y. Wang, R.B. Hayes, S. Yin, N. Titenko-Holland, M. Dosemeci, Y.Z. Wang, P. Kolachana, W. Lu, L. Xi, G.L. Li, M.T. Smith, Benzene increases aneuploidy in the lymphocytes of exposed workers: a comparison of data obtained by fluorescence in situ hybridization in interphase and metaphase cells, Environ. Mol. Mutagen. 34 (1999) 260–268.
50. A. Zijno, P. Leopardi, F. Marcon, R. Crebelli, Analysis of chromosome segregation by means of fluorescence in situ hybridisation: application to cytokinesis-blocked human lymphocytes, Mutat. Res. 372 (1996) 211–219.
51. M.J. Ramirez, J. Surrallés, P. Galofre, A. Creus, R. Marcos, Radioactive iodine induces clastogenic and age-dependent aneugenic effects in lymphocytes of thyroid cancer patients as revealed by interphase FISH, Mutagenesis 12 (1997) 449–455.
52. A. Carere, A. Antoccia, D. Cimini, R. Crebelli, F. Degrassi, P. Leopardi et al., Genetic effects of petroleum fuels: II. Analysis of chromosome loss and hyperploidy in peripheral lymphocytes of gasoline station attendants, Environ. Mol. Mutagen. 32 (1998) 130–138.
53. A. Carere, A. Antoccia, D. Cimini, R. Crebelli, F. Degrassi, P. Leopardi et al., Analysis of chromosome loss and nondisjunction in cytokinesis-blocked lymphocytes of 24 male subjects, Mutagenesis 14 (1999) 491–496.
54. J. Catalán, J. Surrallés, G.C.-M. Falck, K. Autio, H. Norppa, Segregation of sex chromosomes in human lymphocytes, Mutagenesis 15 (2000) 251–255.
55. J. Surrallés, F. Darroudi, A.T. Natarajan, Evidence for low level of DNA repair in human chromosome 1 heterochromatin, Genes Chromosomes Cancer, 20 (1997) 173–184.
56. R. Tice, E.L. Schneider, J.M. Rary, The utilization of bromodeoxyuridine incorporation into DNA for the analysis of cellular kinetics, Exp. Cell Res. 102 (1976) 232–236.
57. R.R. Tice, P. Thorne, E.L. Schneider, BISACK analysis of the phytohaemagglutinin-induced proliferation of human peripheral lymphocytes, Cell Tiss. Kinet. 12 (1979) 1–9.
58. V. Palma, H. Tudon, L. Buentello, S. Nava, P. Ostrosky, F. Salamanca, Methods for the analysis of cellular kinetics in PHA-stimulated blood lymphocytes using BrdU incorporation. A comparative study, Mutat. Res. 286 (1993) 267–273.
59. E.B. Whorton, Some experimental design and analysis considerations for cytogenetic studies, Environ. Mutagen. 7 (Suppl. 4) (1985) 9–15.
60. C. Bolognesi, A. Abbondandolo, R. Barale, R. Casalone, L. Dalpra, M. De Ferrari et al., Age-related increase of baseline frequencies of sister chromatid exchanges, chromosome aberrations, and micronuclei in human lymphocytes, Cancer Epidemiol. Biomarkers Prev. 6 (1997) 249–256.
61. S. Bonassi, M. Ceppi, V. Fontana, F. Merlo, Multiple regression analysis of cytogenetic human data, Mutat. Res. 313 (1994) 69–80.

8. Micronuclei

8.1. Introduction

8.1.1. Purpose of the guideline

To provide concise guidance for evaluating the frequency of micronucleated blood lymphocytes and exfoliated epithelial cells sampled from individuals potentially exposed to clastogens or aneugens. This guideline must be read in conjunction with the General Introduction (Section 2), which contains information on such issues as study design, ethics, and appropriate statistical considerations.

8.1.2. Principles of the assay

Micronuclei (MN) arise in mitotic cells from chromosomal fragments or chromosomes that lag behind in anaphase and are not integrated into the daughter nuclei. Micronuclei harbouring chromosomal fragments result, e.g., from: (A) direct DNA breakage; (B) replication on a damaged DNA template; and (C) inhibition of DNA synthesis (see Guideline on Chromosomal Aberrations). MN harbouring whole chromosomes are primarily formed from failure of the mitotic spindle, kinetochore, or other parts of the mitotic apparatus or by damage to chromosomal sub-structures, alterations in cellular physiology, and mechanical disruption. Thus, an increased frequency of micronucleated cells is a biomarker of genotoxic effects that can reflect exposure to agents with clastogenic (chromosome breaking; DNA as target) or aneugenic (aneuploidogenic; effect on chromosome number; mostly non-DNA target) modes of action.

In human population studies, the frequency of MN is most commonly determined in cultured peripheral blood lymphocytes after being stimulated to proliferate by phytohaemagglutinin (PHA). The most frequently applied methodology uses the cytokinesis-block micronucleus technique [1] in which scoring is limited to cells that have divided once since mitogen-stimulation. MN observed in cultured lymphocytes are believed to arise primarily in vitro from (A) chromatid-type chromosomal aberrations formed during DNA replication on a damaged template, (B) chromosome-type aberrations initiated before the mitosis and duplicated at replication, or (C) disturbances of mitotic apparatus leading to chromosome lagging.

MN arising in vivo, inducible by both clastogenic and aneugenic mechanisms, can be scored in exfoliated epithelial cells [2] sampled, e.g., from buccal or nasal mucosa or urine, or in peripheral blood mononuclear cells (e.g., isolated lymphocytes) [3,4]. Also, blood erythrocytes can be examined for MN, although this approach has been reported to be applicable only to splenectomised individuals in whom micronucleated erythrocytes are not screened from blood circulation [5]. There are indications for the potential use of the erythrocyte MN assay also in non-splenectomised individuals [4], but this approach is not considered in this guideline (see Ref. [6]).

Importantly, the mechanistic origin of individual MN can be determined. MN arising from lagging chromosomes can be identified by the presence of a kinetochore using antikinetochore antibodies or by the presence of centromeric DNA sequences using fluorescence in situ hybridisation (FISH) [7,8]. MN, which do not contain kinetochore/centromeric DNA sequences, are interpreted to harbour acentric chromosomal fragments.

The tissue to be examined should be chosen on the basis of existing knowledge on the action of the exposing agent. The possible or proximate target tissue can be sampled, e.g., when the exposure is expected to act on upper airways (i.e., nasal or buccal cells for suspected nasal carcinogens) or the bladder (cells in urine for suspected bladder carcinogens). Peripheral lymphocytes are used as surrogate target cells.

8.1.3. Definitions

MN are small, extranuclear bodies that are formed in mitosis from acentric chromosomal fragments or chromosomes that are not included in either daughter nucleus. Thus, MN contain either chromosomal fragments or whole chromosomes.

Kinetochore-positive/centromere-positive MN contain whole chromosomes or centric fragments; centric fragments are expected to be rare in MN but may possibly be identified by additional hybridisation with telomeric DNA sequences.

Kinetochore-negative/centromere-negative MN are MN that are interpreted to contain acentric chromosomal fragments.

8.1.4. Significance of the endpoint and application in risk assessment

There is experimental and epidemiological evidence for the association of specific structural and numerical chromosomal aberrations in carcinogenesis [9,10]. Also, at the group level, an increased frequency of structural chromosomal aberrations in peripheral blood lymphocytes is associated with an increased overall risk for cancer [11–13]. No information is available on the possible association of exposure to aneugenic agents and an increased cancer risk. A prospective epidemiological investigation indicated no cancer predictive value for a high level of MN in cultured human lymphocytes. However, due to the considerable methodological variation among the studies included in this evaluation and the relatively limited number of observations, a definite conclusion on the possible association between micronucleus frequency and cancer risk was not possible [11,12].

It is unclear whether MN formation as such has a specific role in carcinogenesis. The process responsible for the formation of MN could be an important mechanism by which cells with a loss of heterozygosity at key genetic loci are produced. Furthermore, MN may reflect genomic instability [14].

8.2. Population characteristics

Important characteristics to consider for the exposed and referent populations are discussed in Section 2. As presented in this section, specific information on socio-economic status and lifestyle (e.g., smoking, recreational/medical drug usage), health status, and endogenous factors (e.g., sex, age) need to be obtained for each subject. Sex and age have clearly been shown to affect MN frequency in cultured lymphocytes [15–28]. These effects are mainly due to MN containing sex chromosomes [24–30]. There is information that low vitamin B_{12}/folate and high homocysteine status are associated with an increased frequency of MN [31–33]. Thus, such information, whenever available, should be taken into account when evaluating the results.

When feasible, individuals with genetic polymorphisms associated with increased susceptibility to genotoxic agents can be identified. Although no consistent association between genotypes and MN frequency has emerged in human population studies available [34–41] such modulation would not be unexpected based on the ability of *GSTM1* and *NAT2* genotypes to modulate the frequency of chromosomal aberrations [42–48]. A concurrent referent population matched for sex and age with the exposed group is necessary.

As discussed in Section 2, the number of subjects needed for a particular study design depends, for a desired power, on the magnitude of the expected effect and the known variability of the measured endpoint. Information on optimal population size characteristics for MN in human blood lymphocytes or other cell types should always be considered during study design. Appropriate software programs are available to assist the investigator in making these calculations (see Section 2). Information on population size calculations for MN are not currently available.

Where the number of subjects to be evaluated cannot be determined prior to study implementation or in situations where the appropriate numbers of subjects are not available, the uncertainty of the observed effect should be considered in evaluating the results. As many factors are known to influence baseline MN level, exposure effects may not be expected in studies on few individuals.

8.3. Sample timing

The optimal sample time for any cell population is during a long-term chronic exposure when the induction and repair of DNA damage (or aneugenic damage) is presumed to be at steady state. This sample time maximises the likelihood of identifying an exposure to a clastogenic or aneugenic agent. For other exposure conditions, the optimal sample time depends to some extent on the cell type of interest. In the ex vivo lymphocyte method, it is particularly important that the exposed and referent populations are sampled concurrently, to minimise technical variation associated with cell culturing. Strictly concurrent sampling is not necessary for direct cell samples, but even here the interval between samplings should be minimised whenever possible to avoid possible seasonal effects on MN frequencies or fluctuation in cell preparation techniques.

8.3.1. Cultured lymphocytes

Following an acute exposure or after termination of a chronic exposure to an S-dependent clastogen, the optimal collection time for detecting induced damage capable of being expressed as MN ex vivo probably depends on the DNA lesions but can be expected to be within a few hours of exposure termination. The frequency of DNA damage in lymphocytes may decrease with increasing time between exposure and sampling due to DNA repair processes, the loss of heavily damaged cells from the blood through apoptosis or necrosis, and cell turnover. The effect of exposures inducing chromosome-type aberrations is expected to be detectable for a longer period. In general, it may be recommended that blood samples are collected within 2 days after the exposure; they may be collected at a later time but with an increasing likelihood that a positive MN effect will not be detectable. If the DNA lesions induced are slowly removed, an effect may be detectable several months after the exposure. Detailed follow-up studies using the MN assay are, however, few and do not allow accurate estimation of lesion half-lives. It is presently unclear whether exposure of resting blood lymphocytes to aneugenic agents, primarily affecting dividing cells, could be detectable by MN analysis of cultured lymphocytes.

8.3.2. Uncultured lymphocytes

In uncultured blood lymphocytes, MN most likely reflect clastogenic and aneugenic events that occurred during the last in vivo cell division. As that division may have occurred anytime from within several hours to several years earlier, depending on the lifetime of the lymphocyte subset studied, this cell population may not be useful in the case of acute or relatively short exposure situations, but might be applicable in long-term exposures. Information on optimum sampling times is, however, presently not available.

8.3.3. Epithelial cells

In epithelial cells, the critical exposure period for MN formation is during the last division prior to sampling. As the exact duration between the last division in the basal layer and migration to the surface is not usually known, it may not be possible to estimate the correct sampling time following acute and short exposures. Thus, because the in vivo epithelial cell MN assay reflects effects in a rather narrow exposure window, it is particularly applicable in situations where the clastogenic or aneugenic exposure under study has been long-term.

8.4. Methods

Descriptions of the cytokinesis-block MN method [1,16,49] and other MN techniques (e.g., Refs. [3,7,8,32,33,49–55]) are available.

8.4.1. Quality assurance:

A system for recording the collection, transport, and storage of samples should be established to ensure that sample integrity is maintained.

The use of coded samples where scoring is conducted without knowledge of exposure information is critical for eliminating potential bias.

Although not often employed in human population studies, it is important to include a quality assurance step to demonstrate the adequacy of the criteria used to detect an increase in MN. In some situations, subsets of the population (e.g., women versus men; aged versus young) may serve, retrospectively, to demonstrate the adequacy of the criteria used during scoring. However, this approach may not always be successful, as the groups studied may mainly or exclusively consist of one sex or have a narrow age distribution. It would be more appropriate to identify and specifically include, as part of the study design, an internal positive control among the samples being analysed. For the ex vivo method with lymphocytes, this can be accomplished by including fixed, stored cells or previously prepared slides with a known induced frequency of MN in the study design. This approach assesses the ability of the scorers to appropriately detect micronuclei and must be accomplished in such a manner that the identity of the samples/slides remains unknown to the scorer. In all cases, an appropriate concurrent negative control against which to compare the positive control response, should be included.

In addition to these internal quality assurance steps, inclusion of external quality assurance steps should be considered in the study design. External steps could involve the sharing of replicate samples with an independent laboratory or the independent

re-scoring by another laboratory of a certain fraction of the total slides scored.

8.4.2. Tissue collection

Because of the risk for infectious diseases (hepatitis, HIV), appropriate precautions must always be followed when handling tissue samples (see Ref. [56,57]).

When blood is used, it is mixed with heparin at sampling. It is recommended that blood samples be processed within 24 h of collection but this may not always be possible, and lymphocytes from blood stored for several days can be cultured successfully. However, especially for S-dependent clastogenic agents, increasing the time between collection and processing is likely to be accompanied by a decline in DNA damage due to DNA repair. To minimise loss of DNA damage due to DNA repair processes, it would be useful to refrigerate the blood samples at below 8°C between collection and processing. In many situations, blood collection will occur at sites distant from the laboratory and transportation will be necessary. If blood samples are shipped commercially, a minimum–maximum thermometer in the shipping container will provide a measure of the temperature range experienced. If air transportation is used, the blood should not be subjected to X-irradiation at security checkpoints. A sheet of X-ray film can be included in the shipping package to verify if such an exposure occurred, although the doses involved in this type of X-irradiation are low and probably not detectable by MN analysis. If samples are shipped internationally, appropriate permits should be obtained in advance and included in the shipment to avoid delays at customs. Most importantly, blood from both exposed and non-exposed individuals must be handled in the same manner. Where blood collection is protracted across time, a balanced collection design must be used.

If uncultured cells are being studied, they are usually processed to microscopic slides immediately or shortly after the sampling, to assure that the specimens are of high quality.

8.4.3. Description

8.4.3.1. Direct samples. Cells are collected by scraping with a spatula, toothbrush, tissue sample brush, or by other appropriate means, from the desired location (e.g., buccal or nasal mucosa) [50]; from urine, cells can be collected by centrifugation [51]. Lymphocytes may be identified in smear preparations [4] or a specific subset of lymphocytes may be isolated from a blood sample by appropriate methods such as immunomagnetic separation [3]. For possible sources of accessible cells, see Ref. [2].

The cells are either directly smeared on slides, or suspended in an appropriate buffer solution and applied on slides after centrifugation, manually or by cytocentrifugation. Especially with lymphocytes, the technique may involve a step where hypotonic treatment is used to enlarge the cells.

8.4.3.2. Lymphocyte cultures. Selection of culture medium, serum type and concentration, and the use of whole-blood versus purified lymphocyte cultures should be based on experience. The technique used should provide good lymphocyte growth. It is recommended that the same lot of media, serum, and other reagents be used, whenever possible, throughout a study. For cultures of isolated lymphocyte cultures, it is necessary to determine the total white cell count and to prepare a slide for a differential count from each blood sample, to standardise the number of lymphocytes in the cultures.

Whole blood or isolated mononuclear cells are added to culture medium containing PHA and incubated at 37°C. If the culture containers are not sealed tightly, the cultures should be maintained in a humidified incubator equilibrated with 5% CO_2. At least duplicate cultures are established for each blood sample, but more cultures may be established if deemed useful. To limit MN analysis to cells that have divided once in vitro, cytochalasin B (Cyt-B) is usually added to the cell cultures (cytokinesis-block micronucleus technique) [1,48]. Cyt-B blocks cytokinesis (which normally occurs in telophase); cells that have divided once in vitro can be identified by the presence of two nuclei (binucleate cells). If the proliferation status of the lymphocytes scored is not identified, the MN represent a mixture of in vivo and in vitro events which cannot be distinguished from each other. Depending on the proportion of MN produced in vitro, MN frequency in such cells may reflect, in addition to genotoxic exposure, mitotic activity of the cultured cells [1]. Although the cytoki-

nesis-block MN technique is the method of choice in lymphocyte MN studies, cycling cells can also be identified by labelling replicating DNA with bromodeoxyuridine (BrdU) [58] and detecting the labelled cells using a monoclonal BrdU-DNA antibody [59]. As also some undivided cells in S- or G2-stage of cell cycle and cells that have completed more than one division may be labelled with BrdU, the method appears less accurate than the cytokinesis block technique in identifying the correct cell population. By appropriate timing of BrdU labelling and cell culture, it might be possible to label primarily cells that have completed their first in vitro division [41], but this approach will have to be further evaluated. It is also unclear if BrdU affects MN frequencies. If the cytokinesis-block technique is used, Cyt-B is added at an appropriate time, to block cytokinesis of most cells in their first in vitro division. The correct timing of Cyt-B addition depends on the cell culture system used. In most situations, addition of Cyt-B at 24 h of culture assures that cells are blocked in their first in vitro division, but later time points (e.g., 44 h) can be used if a culture technique with slow lymphocyte growth is used. The concentration of Cyt-B should be one demonstrated to maximise the frequency of binucleate cells without affecting nuclear division or increasing the background frequency of MN (usually 3–6 μg/ml) [60]. The cells are harvested at the shortest culture time that provides enough cells for analysis. In the cytokinesis-block method, the culture time should be adjusted so that the proportion of multinucleated cells (cells with three or more nuclei) is as low as possible. Each culture is harvested and processed separately. Treatment with a hypotonic solution may be used to swell the cells for more accurate scoring. The cells are then either smeared or cytocentrifuged on slides and fixed, or (customary for whole-blood cultures) fixed in suspension with methanol–acetic acid prior to slide preparation. The preservation of cytoplasm is important for identifying binucleate cells.

8.4.4. Slide storage and staining

Before analysis, the microscopic slides or fixed cell suspensions should be stored in a manner that maintains their high quality. Depending on the type of the samples, storage may occur in a freezer (−20°C; methanol–acetic acid fixed cell suspensions, slides for FISH) or at room temperature (mounted or unmounted fixed slides, slides in methanol).

The slides are stained by appropriate dyes for transmitted-light (e.g., Giemsa) or fluorescence (e.g., acridine orange, DAPI, Hoechst 33342) microscopy. If the mechanistic origin of MN are identified by FISH using pancentromeric DNA probes, antikinetochore antibodies, or other molecular cytogenetic methods, the techniques used should provide definitive labelling of the centromere/kinetochore area of all chromosomes.

8.4.5. MN analysis

Uniform criteria for scoring MN, such as those contained in Ref. [50,61,62] should be used. Only MN not exceeding 1/3 of the main nucleus diameter, clearly separable from the main nucleus and with distinct borders and of the same colour as the nucleus, should be scored. Where multiple scorers are involved, a balanced scoring design must be used. Scoring should be conducted by a trained and experienced observer.

Slides must be scored without knowledge of the exposure group; slides from the quality assurance standard should be included within this set of slides. Where possible, the number of cells to be analysed per subject should be based upon the frequency of cells with damage in the referent population, or on expectations based on historical control data. In general, the goal is to obtain as few subjects as possible with zero class data while avoiding unnecessary costs. In practice, 1000–2000 cells are often scored per subject in lymphocyte studies utilising the cytokinesis block MN technique, while more cells (3000–5000 per subject) are evaluated in epithelial cells due to the lower baseline MN frequency. Care should be taken not to include apoptotic cells in the analysis.

In the cytokinesis block MN method, only binucleate cells should be analysed for MN; further divisions of a binucleate cell, usually resulting in cells with 3–4 nuclei, are highly irregular and show high MN rates [8,63]. Irregularly shaped binucleate cells should be disregarded, as these may represent products of further divisions of binucleate cells.

Mononucleate cells represent a heterogeneous group of cells that have either not divided in vitro or escaped the cytokinesis block after one or more divisions, and are, therefore, not scored.

In the cytokinesis block method, the relative proportions of binucleate and multinucleate cells (with three or more nuclei) may be used as a measure of cellular growth [64]. The easiest way of expressing this is to use the cytokinesis-block proliferation index (CBPI [65], which gives the average number of cell divisions completed by the cells scored. The CBPI = [number binucleate cells + 2 (number multinucleated cells)]/(total number of cells).

When molecular cytogenetic techniques are used to identify the mechanistic origin of MN [42–45], the frequency of MN and the presence of a centromere or a kinetochore in a MN should be judged only in those micronucleated cells where the main nucleus shows the appropriate labelling pattern. Alternatively, all cells (also those lacking proper centromere or kinetochore staining) can be scored for the frequency of MN, which, together with the proportion of centromere- or kinetochore-positive MN, is used in calculating the frequencies of MN harbouring whole chromosomes (signal positive MN) and fragments (signal negative MN).

8.4.5.1. Data recording. The number of micronucleated cells among the cells scored and the number of MN in each micronucleated cell are recorded. To facilitate quality assurance, it is a good practice to record microscope co-ordinates for each cell containing MN when using permanent preparations, or to store the images of micronucleated cells in the memory of an imaging station when non-permanent slides (fluorescence techniques) are used.

8.4.6. Sources of assay variation

In the ex vivo assay with lymphocytes, the type and source of culture medium and mitogen used can affect cell cycle progression, MN frequency, and cytokinesis-block replication index. This is a problem especially if Cyt-B is not used, since the proportions of cells that have divided once or more often in vitro or have not divided at all vary among individuals, which may affect MN frequencies.

The type of mitogen influences the proliferation rate of lymphocytes. The stability of the temperature control of incubators, and the temperature variation within the incubators can affect differentially the growth of cultures. In the cytokinesis-block technique, only binucleate cells (which have completed one nuclear division only) should be scored for micronuclei, to avoid bias caused by further division cells and undivided cells. An inadequate concentration of Cyt-B will result in non-optimal cytokinesis block, which may affect MN frequencies [60].

Long culture times result in an increase in lymphocyte MN frequencies, possibly as a reflection of the delayed division of more heavily damaged cells. This may influence MN frequencies since there may be differences among individuals in cell proliferation rates [55].

With all MN assays, the quality of slides and the staining used influence the efficiency of MN detection and thus MN frequency. Therefore, it is important that the slides are of good quality and similar for all samples and that only one valid method is used throughout for slide preparation and staining the slides. The preservation of cytoplasm is important especially for cells fixed in suspension, to avoid loss of MN. In the cytokinesis block technique, cytoplasm boundaries should be clearly visible to facilitate correct identification of binucleate cells.

Variation among scorers is important in all microscopy, including MN analyses. Variation can be decreased by having only one scorer, and, when several scorers are used, reduced by ensuring a balanced scoring design. Thus, for example, each scorer should analyse the same number of cells from the slides of all subjects (the same cells not being scored twice) rather than one scorer analysing all cells from some subjects and another scorer all cells from other subjects. Variation can also be reduced by ensuring that the scoring is not unnecessarily protracted across time. Each scorer should be experienced so that scoring criteria do not evolve during the analysis. An adequate number of cells should be scored from each subject (see Section 8.4.5).

It is presently unclear if the use of Cyt-B affects MN frequencies. On one hand, MN frequency in binucleate cells has been reported not to increase with the concentration of Cyt-B [1], but a dose–response may not either be expected when ''affected cells'' such as binucleates are scored exclusively. On the other hand, inadequate concentrations of Cyt-B

have been reported to increase MN [60]. Furthermore, binucleate cells have been reported to have less MN (per nuclei) and different chromosomal content in MN than cells cultured without Cyt-B [3,28,55]; the reasons are not well understood. A lack of uniform collection, shipping, processing, and storage of samples will increase technical variability. If cells are stored in a fixative solution, the conditions should be the same for all samples.

8.4.7. Statistical analysis

As described in Section 2, data should be evaluated according to the statistical plan of the study. For each subject, the frequency of micronucleated cells or frequency of MN per 1000 cells is calculated. If the whole chromosome content of MN has been determined, these parameters are separately calculated for centromere/kinetochore positive and negative MN. There is no unified agreement on the proper statistical methods to use. However, the unit of exposure is the individual and not the cell; pooling cells across individuals within a study group eliminates a critically important source of variability. The choice of statistical method to be used may depend on the distribution of the data. If the distribution of the data is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, parametric methods based upon normal distribution can also be used, when appropriate, such as after relevant data transformation. When needed, zero values can be transformed [66]. Statistical advice may be needed to determine whether parametric or non-parametric methods are more appropriate for a particular study. The relative and absolute effects of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion by estimating the precision of the effects seen in the study. Stratified analyses can be used for exploring the effect modification by factors such as age, sex, smoking, other exposures, etc. More sophisticated multivariate methods may be necessary for investigating the effect of factors that either may modify or be confounded with the effects [20,67]. Wherever possible, the analysis should be based on individual exposure information. Consideration for the effect of multiple comparisons on the false positive rate should be included in the statistical analysis.

The quality assurance standard is compared against the appropriate concurrent control cultures using appropriate statistics.

8.4.8. Interpretation of results

Criteria for identifying a positive response (i.e., increased frequency of micronucleated cells in an exposed group compared to the referent population; a dose response relationship if variable exposure levels have been established) must be identified prior to implementing the study.

A statistically significantly higher frequency of micronucleated cells among the exposed population in comparison to the referent population is considered a positive result (i.e., the agent of interest induced clastogenic or aneugenic damage in cells of exposed individuals). Presentation of the results in terms of the size of the effect with confidence intervals may also help in the interpretation. A positive result is supported by a dose–response relationship when individuals with different exposure levels are considered. A dose–response relationship cannot, however, be a pre-requisite; an increased MN frequency, especially in peripheral lymphocytes, may reflect an effect of a long-term exposure for which accurate dose estimation is difficult. Maximum confidence that the exposure results in an increased frequency of MN requires reproducible results in independent studies. It is appreciated that independent assessments are not always possible (e.g., in the case of accidents).

Positive results from the ex vivo MN assay indicate that the exposure studied induced genotoxic lesions that were detected as MN when the cells were cultured. MN observed in binucleate lymphocytes obtained using the cytokinesis-block method have, by definition, mostly been formed in vitro (although some in vivo MN may have survived the division) from pre-existing lesions in the cells. For epithelial cells, the turnover of the tissue is continuous, and a positive finding is expected to have resulted from exposure occurring during a relatively short period before the sampling; individual differences in cell proliferation rates may affect MN frequencies.

Information on the presence of whole chromosomes and fragments in MN can be used to identify the nature of the damage induced (i.e., clastogenic or aneugenic).

The lack of a statistically significant increase in micronucleated cells indicates that under the exposure conditions evaluated and for the calculated power of the study, the exposure did not result in a significant increase in chromosomal damage in the cell population evaluated.

The results of the quality assurance control are used to demonstrate the adequacy of the methodology and the scorers in identifying chromosomal damage.

8.5. Report

All data should be presented in tabular and/or graphical form, and include all observed results (see below). This will allow comparisons between laboratories, re-analyses of results by outside reviewers, and provide a database for baseline background values and for exposures to physical and chemical agents. All results of the study should be reported with none omitted, and the definition and handling of ''outliers'' described.

The study report should include the following minimal information:

- copy of the questionnaire used to determine personal characteristics
- statement of ethical approval
- characteristics of the exposed and referent populations
- exposure information and analytical methods used
- tissue collection procedures, including timing in relationship to exposure, transportation, storage, and chain-of-custody
- method for coding samples prior to analysis
- assay conditions for cell cultures: number of cultures and replicates, composition of medium, CO_2 concentration, incubation duration, concentration and treatment schedule for Cyt-B (or BrdU)
- description of the techniques of slide preparations and staining and the principles of slide coding (''blind analysis'')
- number of cells scored per culture and number of replicates
- number of cells analysed (data given separately for each subject)
- method for evaluating mechanistic origin of MN
- criteria for scoring MN and mechanistic origin (if conducted)
- criteria for cell kinetic analysis (if conducted)
- number of scorers used and the scoring design
- kinetic data for each subject (if scored)
- tabulated results for each subject
- statistical methods
- discussion and interpretation of results
- internal and external quality assurance procedures.

8.6. Literature cited

1. M. Fenech, A. Morley, Measurement of micronuclei in lymphocytes, Mutat. Res. 147 (1985) 29–36.
2. S.A. Salama, M. Serrana, W.A. Au, Biomonitoring using accessible human cells for exposure and health assessment, Mutat. Res. 436 (1999) 99–112.
3. J. Surrallés, G. Falck, H. Norppa, In vivo cytogenetic damage revealed by FISH analysis of micronuclei in uncultured human T lymphocytes, Cytogenet. Cell Genet. 75 (1996) 151–154.
4. M. Fenech, G. Perepetskaya, L. Mikhalevich, A more comprehensive application of the micronucleus technique for biomonitoring of genetic damage rates in human populations — Experiences from the Chernobyl catastrophe, Environ. Mol. Mutagen. 30 (1997) 112–118.
5. J.T. MacGregor, C.M. Wehr, R. Hiatt, B. Peters, J.D. Tucker, R.G. Langlois et al., ''Spontaneous'' genetic damage in man: evaluation of interindividual variability, relationship among markers of damage, and influence of nutritional status, Mutat. Res. 377 (1997) 125–135.
6. The International Collaborative Project on Micronucleus Frequency in Human Populations (HUMN). http://ehs.sph.berkeley.edu/holland/HUMN/humn.html. Contact Michael Fenech, Chairman, Coordinating Committee (michael.fenech@dhn.csiro.au).

7. E.J. Thomson, P.E. Perry, The identification of micronucleated chromosomes: a possible assay for aneuploidy, Mutagenesis 3 (1988) 415–418.
8. H. Norppa, L. Renzi, C. Lindholm, Detection of whole chromosomes in micronuclei of cytokinesis-blocked human lymphocytes by antikinetochore staining and in situ hybridization, Mutagenesis 8 (1993) 519–525.
9. F. Mitelman, Chromosomes, genes, and cancer [editorial], CA Cancer J. Clin. 44 (1994) 133–135.
10. D.A. Rew, Significance of aneuploidy, Br. J. Surg. 81 (1994) 1416–1422.
11. L. Hagmar, A. Brøgger, I.-L. Hansteen, Heim S, B. Högstedt, L. Knudsen et al., Cancer risk in humans predicted by increased levels of chromosomal aberrations in lymphocytes: Nordic Study Group on the Health Risk of Chromosome Damage, Cancer Res. 54 (1994) 2919–2922.
12. L. Hagmar, S. Bonassi, U. Strömberg, A. Brøgger, L.E. Knudsen, H. Norppa, C. Reuterwall, Chromosomal aberrations in lymphocytes predict human cancer: a report from the European Study Group on Cytogenetic Biomarkers and Health (ESCH), Cancer Res. 58 (1998) 4117–4121.
13. S. Bonassi, A. Abbondandolo, L. Camurri, L. Dal Prá, M. De Ferrari, F. Degrassi et al., Are chromosome aberrations in circulating lymphocytes predictive of a future cancer onset in humans? Preliminary results of an Italian cohort study, Cancer Genet. Cytogenet. 79 (1995) 133–135.
14. A. Inoue, K. Yokomori, H. Tanabe, H. Mizusawa, T. Sofuni, Y. Hayashi, Y. Tsuchida, H. Shimatake, Extensive genetic heterogeneity in the neuroblastoma cell line NB(TU)1, Int. J. Cancer 72 (1997) 1070–1077.
15. B. Högstedt, B. Åkesson, K. Axell, B. Gullberg, F. Mitelman, R.W. Pero et al., Increased frequency of lymphocyte micronuclei in workers producing reinforced polyester resin with low exposure to styrene, Scand. J. Work Environ. Health 241 (1983) 186–192.
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. J.W. Yager, M. Sorsa, S. Selvin, Micronuclei in cytokinesis-blocked lymphocytes as an index of occupational exposure to alkylating cytostatic drugs, Methods for Detecting DNA-Damaging Agents in Man: Applications in Cancer Epidemiology and Prevention, International Agency for Research on Cancer, Lyon, France 1988, vol. 89, pp. 213–216.
18. Nordic Study Group on the Health Risk of Chromosome Damage, A Nordic data base on somatic chromosome damage in humans, Mutat. Res. 241 (1990) 325–337.
19. M. Fenech, S. Neville, J. Rinaldi, Sex is an important variable affecting spontaneous micronucleus frequency in cytokinesis-blocked lymphocytes, Mutat. Res. 313 (1994) 203–207.
20. C. Bolognesi, A. Abbondandolo, R. Barale, R. Casalone, L. Dalpra, M. De Ferrari et al., Age-related increase of baseline frequencies of sister chromatid exchanges, chromosome aberrations, and micronuclei in human lymphocytes, Cancer Epidemiol. Biomarkers Prev. 6 (1997) 249–256.
21. R. Barale, L. Chelotti, T. Davini, S. Del Ry, M.G. Andreassi, M. Ballardin et al., Sister chromatid exchange and micronucleus frequency in human lymphocytes of 1,650 subjects in an Italian population: II. Contribution of sex, age, and lifestyle, Environ. Mol. Mutagen. 31 (1998) 228–242.
22. F. Richard, M. Muleris, B. Dutrillaux, The frequency of micronuclei with the X chromosome increases with age in human females, Mutat. Res. 316 (1994) 1–7.
23. J.C. Hando, J. Nath, J.D. Tucker, Sex chromosomes, micronuclei and aging in women, Chromosoma, 103 (1994) 186–192.
24. M. Guttenbach, R. Schakowski, M. Schmid, Aneuploidy and ageing: sex chromosome exclusion into micronuclei, Hum. Genet. 94 (1994) 295–298.
25. J. Catalán, K. Autio, M. Wessman, C. Lindholm, S. Knuutila, M. Sorsa, H. Norppa, Age-associated micronuclei containing centromeres and the X chromosome in lymphocytes of women, Cytogenet. Cell Genet. 68 (1995) 11–16.
26. J. Nath, J.D. Tucker, J.C. Hando, Y chromosome aneuploidy, micronuclei, kinetochores and aging in men, Chromosoma 103 (1995) 725–731.

27. J.C. Hando, J.D. Tucker, M. Davenport, J. Tepperberg, J. Nath, X chromosome inactivation and micronuclei in normal and Turner individuals, Hum. Genet. 100 (1997) 624–628.
28. J. Catalán, K. Autio, E. Kuosma, H. Norppa, Age-dependent inclusion of sex chromosomes in lymphocyte micronuclei of man, Am. J. Hum. Genet. 63 (1998) 1464–1472.
29. M. Fenech, Chromosomal damage rate, aging, and diet, Ann. N. Y. Acad. Sci. 854 (1998) 23–36.
30. M. Fenech, I.E. Dreosti, J.R. Rinaldi, Folate, vitamin B-12, homocysteine status and chromosome damage in lymphocytes of older men, Carcinogenesis 18 (1997) 1329–1336.
31. M. Fenech, C. Aitken, J. Rinaldi, Folate, vitamin B-12, homocyteine status and DNA damage in young Australian adults, Carcinogenesis 19 (1998) 1163–1171.
32. Y. Odagiri, H. Uchida, Influence of serum micronutrients on the incidence of kinetochore-positive or -negative micronuclei in human peripheral blood lymphocytes, Mutat. Res. 415 (1998) 35–45.
33. N. Titenko-Holland, R.A. Jacob, N. Shang, A. Balaraman, M. Smith, Micronuclei in lymphocytes and exfoliated buccal cells of postmenopausal women with dietary changes in folate, Mutat. Res. 417 (1998) 101–114.
34. G. van Poppel, P. Verhagen, P.J. van't Veer, P.J. Van Bladeren, Markers for cytogenetic damage in smokers: association with plasma antioxidants and glutathione *S*-transferase, Cancer Epidemiol. Biomarkers Prev. 2 (1993) 441–447.
35. U. Carstensen, A.K. Alexandrie, B. Högstedt, A. Rannug, I. Bratt, L. Hagmar, B- and T-lymphocyte micronuclei in chimney sweeps with respect to genetic polymorphism for CYP1A1 and GST1 (class Mu), Mutat. Res. 289 (1993) 187–195.
36. M. Ichiba, L. Hagmar, A. Rannug, B. Högstedt, A.-K. Alexandrie, U. Carstensen, K. Hemminki, Aromatic DNA adducts, micronuclei and genetic polymorphism for CYP1A1 and GST1 in chimney sweeps, Carcinogenesis 15 (1994) 1347–1352.
37. T.-J. Cheng, D.C. Christiani, X. Xu, J.C. Wain, J.K. Wiencke, K.T. Kelsey, Increased micronucleus frequency in lymphocytes from smokers with lung cancer, Mutat. Res. 349 (1996) 43–50.
38. R. Scarpato, L. Migliore, A. Hirvonen, G. Falck, H. Norppa, Cytogenetic monitoring of occupational exposure to pesticides: Characterization of GSTM1, GSTT1, and NAT2 genotypes, Environ. Mol. Mutagen. 27 (1996) 263–269.
39. M. Sorsa, S. Osterman-Golkar, K. Peltonen, S.T. Saarikoski, R. Sram, Assessment of exposure to butadiene in the process industry, Toxicology, 113 (1996) 77–83.
40. R. Srám, P. Rössner, K. Peltonen, K. Podrazilova, G. Mrackova, N.A. Demopoulos et al., Chromosomal aberrations, sister-chromatid exchanges, cells with high frequency of SCE, micronuclei and comet assay parameters in 1, 3-butadiene-exposed workers, Mutat. Res. 419 (1998) 145–154.
41. G.-M. Falck, A. Hirvonen, R. Scarpato, S.T. Saarikoski, L. Migliore, H. Norppa, Micronuclei in blood lymphocytes and genetic polymorphism for GSTM1, GSTT1 and NAT2 in pesticide-exposed greenhouse workers, Mutat. Res. 441 (1999) 225–237.
42. N. Conforti-Froes, R. el-Zein, S.Z. Abdel-Rahman, J.B. Zwischenberger, W.W. Au, Predisposing genes and increased chromosome aberrations in lung cancer cigarette smokers, Mutat. Res. 379 (1997) 53–59.
43. R. el-Zein, N. Conforti-Froes, W.W. Au, Interactions between genetic predisposition and environmental toxicants for development of lung cancer, Environ. Mol. Mutagen. 30 (1997) 196–204.
44. H. Norppa, Cytogenetic markers of susceptibility: Influence of polymorphic carcinogen-metabolizing enzymes, Environ. Health Perspect. 105 (Suppl. 4) (1997) 829–835.
45. R. Scarpato, A. Hirvonen, L. Migliore, G. Falck, H. Norppa, Influence of GSTM1 and GSTT1 polymorphisms on the frequency of chromosome aberrations of smokers and pesticide-exposed greenhouse workers, Mutat. Res. 389 (1997) 227–235.
46. S.Z. Abdel-Rahman, R.A. el-Zein, J.B. Zwischenberger, W.W. Au, Association of the NAT1* 10 genotype with increased chromosome

aberrations and higher lung cancer risk in cigarette smokers, Mutat. Res. 398 (1998) 43–54.
47. S. Landi, H. Norppa, G. Frenzilli, G. Cipollini, I. Ponzanelli, R. Barale, A. Hirvonen, Individual sensitivity to cytogenetic effects of 1,2:3,4-diepoxybutane in cultured human lymphocytes: influence of glutathione *S*-transferase M1, P1 and T1 genotypes, Pharmacogenetics, 8 (1998) 461–471.
48. L.E. Knudsen, H. Norppa, M.O. Gamborg, P.S. Nielsen, H. Okkels, H. Soll-Johanning et al., Chromosomal aberrations induced by urban air pollution in humans: influence of DNA repair and polymorphisms of glutathione *S*-transferase M1 and *N*-acetyltransferase 2, Cancer Epidemiol. Biomarkers Prev. 8 (1999) 303–310.
49. M.F. Fenech, The cytokinesis-block micronucleus technique, in: P. Gerd (Ed.), Technologies for Detection of DNA Damage and Mutations, Pfeiffer, Plenum Press, New York, 1996, pp. 25–36.
50. C. Ballarin, F. Sarto, L. Giacomelli, G.B. Bartolucci, E. Clonfero, Micronucleated cells in nasal mucosa of formaldehyde-exposed workers, Mutat. Res. 280 (1992) 1–7.
51. H.F. Stich, W. Stich, M.P. Rosin, The micronucleus test on exfoliated cells, in: A. Muhammed, R.C. von Borstel (Eds.), Basic and Applied Mutagenesis, Plenum, New York, 1985, pp. 337–342.
52. M. Fenech, A.A. Morley, Kinetochore detection in micronuclei: an alternative method for measuring chromosome loss, Mutagenesis 4 (1989) 98–104.
53. N. Titenko-Holland, A.J. Levine, M.T. Smith, P.J.E. Quintana, M. Boeniger, R. Hayes et al., Quantification of epithelial cell micronuclei by fluorescence in situ hybridization (FISH) in mortuary science students exposed to formaldehyde, Mutat. Res. 371 (1996) 237–248.
54. J. Surrallés, K. Autio, L. Nylund, H. Järventaus, H. Norppa, T. Veidebaum et al., Molecular cytogenetic analysis of buccal cells and lymphocytes from benzene exposed workers, Carcinogenesis 18 (1997) 817–823.
55. G. Falck, J. Catalán, H. Norppa, Influence of culture time on the frequency and contents of human lymphocyte micronuclei with and without cytochalasin B, Mutat. Res. 392 (1997) 71–79.
56. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
57. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health. Biosafety in Microbiological and Biomedical Laboratories. HHS Publication No (CDC) 93-8395. 3rd edn. (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H. Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices. 2nd edn., ASM Press, Washington DC, 1995.
58. M. Pincu, D. Bass, A. Norman, An improved micronuclear assay in lymphocytes, Mutat. Res. 139 (1984) 61–65.
59. H. Norppa, M. Hayashi, J. Mäki-Paakkanen, M. Sorsa, The micronucleus assay in lymphocytes, in: M.L. Mendelsohn, R.J. Albertini (Eds.), Mutation and the Environment Part B: Metabolism, testing Methods and Chromosomes. Wiley-Liss, New York, 1990, pp. 207–216.
60. A. Zijno, F. Marcon, P. Leopardi, R. Crebelli, Simultaneous detection of chromosome loss and non-disjunction in cytokinesis blocked human lymphocytes by in situ hybridization with a centromeric DNA probe; implications for the human lymphocyte in vitro micronucleus assay using cytochalasin B, Mutagenesis 9 (1994) 225–232.
61. 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.
62. P.E. Tolbert, C.M. Shy, J.W. Allen, Micronuclei and other abnormalities in buccal smears: methods development, Mutat. Res. 271 (1992) 69–77.
63. C. Lindholm, H. Norppa, M. Hayashi, M. Sorsa, Induction of micronuclei and anaphase aberrations by cytochalasin B in human lymphocyte cultures, Mutat. Res. 260 (1991) 369–375.
64. D. Eastmond, J. Tucker, Identification of aneuploidy-inducing agents using cytokinesis blocked human lymphocytes and an antikinetochore antibody, Environ. Mol. Mutagen. 13 (1989) 34–43.

65. J. Surrallés, N. Xamena, A. Creus, J. Catalán, H. Norppa, R. Marcos, Induction of micronuclei by five pyrethroid insecticides in whole-blood and isolated human lymphocyte cultures, Mutat. Res. 341 (1995) 169–184.
66. E.B. Whorton, Some experimental design and analysis considerations for cytogenetic studies, Environ. Mutagen. 7 (Suppl. 4) (1985) 9–15.
67. S. Bonassi, M. Ceppi, V. Fontana, F. Merlo, Multiple regression analysis of cytogenetic human data, Mutat. Res. 313 (1994) 69–80.

9. Sister chromatid exchanges (SCEs) in lymphocytes

9.1. Introduction

9.1.1. Purpose of the guideline

To provide concise guidance for methodologies used for evaluating the frequency of SCEs in peripheral blood lymphocytes sampled from individuals potentially exposed to genotoxic agents. This guideline must be read in conjunction with the General Introduction (Section 2), which contains information on such issues as study design, ethics, and appropriate statistical considerations.

9.1.2. Principles of the assay

SCEs arise from the reciprocal exchange of DNA between two sister chromatids of a duplicated chromosome (see Refs. [1–3] for a compilation of articles on the nature of SCEs and their utility for human biomonitoring). The frequency of SCEs in eukaryote cells is increased by exposure to genotoxic agents that induce DNA damage (e.g., alkylated bases, crosslinks) capable of interfering with DNA replication; the frequency of SCEs is not increased by genotoxic agents that induce strand breakage only [3]. The currently used method for the detection of SCEs requires DNA replication in the presence of bromodeoxyuridine (BrdU) for two consecutive cell cycles (or at least the first of two consecutive cell cycles), with scoring conducted in second-division metaphase cells. The SCE assay yields quantifiable data from every cell scored, which increases the efficiency of data collection and the identification of DNA damage resulting from exposure to genotoxic carcinogens compared to traditional chromosomal aberration analysis.

Because of the generally ready availability of blood and the need for cycling cells, the focus of this guideline is on mitogen-stimulation blood lymphocytes.

9.1.3. Definitions

SCEs result from the interchange of DNA replication products at apparently identical loci of the sister chromatids of a chromosome in response to a damaged DNA template. The exchange process involves DNA breakage and reunion, although little is known about its molecular basis. SCEs appear to be a consequence of errors of DNA replication on a damaged template, possibly at the replication fork [3–5].

9.1.4. Significance of the endpoint and application in risk assessment

The ready quantifiable nature of SCEs with high sensitivity for revealing toxicant–DNA interaction and the demonstrated ability of genotoxic chemicals to induce a significant increase in SCEs in cultured cells and in cells sampled from treated animals, has resulted in this endpoint being used an indicator of DNA damage in blood lymphocytes of individuals exposed to genotoxic carcinogens (e.g., Refs. [6,7]). However, due to the unclear mechanism of formation and uncertainty of the biological significance of SCEs, this endpoint is not used frequently as a routine test in genetic toxicology. These considerations also apply when SCEs are used as an endpoint in studies of human exposure to genotoxic carcinogens.

No association has been observed between high SCE frequencies and risk of cancer [8,9]. The baseline levels of SCEs fluctuate among individuals and between studies, which makes a general classification of subjects into a high, medium, or low SCE group difficult. Consequently, the possible association of SCE level with cancer risk is more difficult to evaluate than in the case of chromosomal aberrations and needs further evaluation. However, a major part of this background variation may be due to technical rather than true biological variation in the endpoint and could be controlled in a well-conducted study.

9.2. Population characteristics

Important characteristics to consider for the exposed and referent populations are discussed in Section 2. As presented in this section, specific information on socio-economic status and lifestyle (e.g., smoking, recreational/medical drug usage), health status, and endogenous factors, such as age and sex, all of which may affect SCE levels [10–14], need to be obtained for each subject. As SCE frequencies can vary from one study to another, and this variation may be higher than the effect associated with exposure to genotoxic carcinogens, it is very important that blood samples from the referents are collected and processed together with those from the exposed subjects. The glutathione *S*-transferase T1 genotype affects the ''baseline'' frequency of SCE, with GSTT1 null subjects showing higher values [15–18]. This genotype (and other genotypes, if shown to affect SCEs) should be considered when evaluating the results.

As discussed in Section 2, the number of subjects needed for a particular study design depends, for a desired power, on the magnitude of the expected effect, and the known variability of the measured endpoint. Information on optimal population size characteristics for SCEs in human blood lymphocytes should always be considered during study design. Appropriate software programs are available to assist the investigator in making these calculations (see Section 2). Information on population size calculations for SCE can be found also in Ref. [19]. As the effect of genotoxic exposure on SCE frequency is usually expected to be relatively small (e.g., a 10–20% increase by smoking) in comparison with the background level of SCEs, effects may not be expected in studies on few individuals.

Where the number of subjects cannot be determined prior to study implementation or in situations where the appropriate number of subjects are not available, the uncertainty of the observed effect should be considered in evaluating the results.

9.3. Sample timing

As the SCEs detected in cultured peripheral lymphocytes are formed during the S-phase of the cell cycle from pre-existing DNA lesions, the optimal sample time for peripheral blood lymphocytes is during a long-term chronic exposure when the induction and repair of DNA damage in lymphocytes is presumed to be at steady-state. For the sampling of blood after an acute exposure or after termination of a chronic exposure to a genotoxic agent, the optimal collection time for detecting induced DNA damage is within a few hours of exposure termination. The frequency of induced SCEs is expected to decrease with increasing time between exposure and sampling, due to the repair of DNA damage, the loss of heavily damaged cells from the blood through apoptosis or necrosis, and to cell turnover. There are, however, some long-lived T lymphocytes where damage may persist. In general, blood samples should be collected within 2 days after the exposure; they may be collected at a later time but with an increasing likelihood that a positive SCE effect will not be detectable.

9.4. Methods

9.4.1. Quality assurance

A system for recording the collection, transport, and storage of samples should be established to ensure that sample integrity is maintained.

The use of coded samples where scoring is conducted without knowledge of exposure information is critical for eliminating potential bias.

Although seldom employed in past human population studies, it is important to include a quality assurance step to demonstrate the adequacy of the methodology used in the study. In some situations, subsets of the population (e.g., smokers versus nonsmokers) may serve, retrospectively, to demonstrate the adequacy of the methodology. However, this approach may not always be successful, as the study may have too few or no heavy smokers. It would be more appropriate to include, as part of the study design, a quality assurance control among the samples being analysed. This can be accomplished, for example, by including fixed, stored cells or previously prepared slides with a known induced frequency of SCE in the study design. This approach assesses the ability of the scorers to appropriately detect SCE and must be accomplished in such a

manner that the identity of the samples/slides remains unknown to the scorer.

In addition to these internal quality assurance steps, inclusion of external quality assurance steps should be considered in the study design. External steps could involve the sharing of replicate blood samples with an independent laboratory or the independent re-scoring by another laboratory of a certain fraction of the total slides scored.

In all cases, an appropriate concurrent negative control against which to compare the positive control response should be included.

9.4.2. Blood collection

Because of the risk for infectious diseases (hepatitis, HIV), appropriate precautions must always be followed when handling blood samples (see Refs. [20,21]).

At the time of collection, blood is mixed with heparin to prevent coagulation. It is recommended that blood samples be processed within 24 h of collection but this may not always be possible and lymphocyte from blood stored for at least several days have been cultured successfully. However, increasing the time between collection and processing is likely to be accompanied by a decline in DNA damage, due to DNA repair. To minimise the removal of DNA damage by DNA repair processes, it would be useful to refrigerate the blood samples at below 8°C between collection and processing. In many situations, blood collection will occur at sites distant from the laboratory and transportation will be necessary. If blood samples are shipped commercially, a minimum–maximum thermometer in the shipping container will provide a measure of the temperature range experienced. If air transportation is used, the blood should not be subjected to X-irradiation at security checkpoints. If considered possible, a sheet of X-ray film can be included in the shipping package to verify that such an exposure occurred and the information considered during interpretation of the results. If samples are shipped internationally, appropriate permits should be obtained in advance and included in the shipment to avoid delays at customs. However, these doses are very low and probably not detectable using cytogenetics.

Most importantly, blood from both exposed and non-exposed individuals must be handled in the same manner. Where blood collection is protracted across time, a balanced collection design must be used.

9.4.3. Description

9.4.3.1. Culturing. It may be useful to determine the total white cell count and to prepare a slide for a differential count from each blood sample just prior to establishing the cultures. These endpoints may be useful during data interpretation. Selection of culture medium, serum type and concentration, and the use of whole blood versus purified lymphocyte cultures are arbitrary and should be based on experience. It is recommended that the same batch of media, serum, and other reagents be used, whenever possible, throughout a study.

Whole blood or isolated mononuclear cells are added to culture medium containing phytohaemagglutinin (PHA) (if sufficiently justified, other mitogens may be used) and BrdU, and incubated at 37°C. If the culture containers are not sealed tightly, the cultures should be maintained in a humidified incubator equilibrated with 5% CO_2. The cultures should be carefully shielded from light in all steps until cell harvest, to avoid photolysis of BrdU-substituted DNA. The concentration of BrdU used is critical and should be one demonstrated not to affect cell kinetics (typically 10–20 μM). BrdU must be added in the beginning of the culture or no later than 24 h after the addition of the mitogen, PHA, to avoid obtaining cells that have undergone partial S phases giving a confusing staining pattern. The accuracy of BrdU solutions used and the concentration of BrdU in the medium are important, because BrdU influences SCE level. At least duplicate cultures are established for each blood sample, but more cultures may be established if deemed useful (see Ref. [22]).

9.4.3.2. Termination. Incorporation of BrdU into replicating cells allows for the unequivocal identification of second division metaphase cells. The culture time is chosen according to the culture system used and should result in adequate proportion of second division metaphase cells, to facilitate rapid analysis.

Cell cultures are treated with colchicine or Colcemid (generally one to two h prior to harvesting. Culture are harvested and processed through hypotonic (e.g., 0.075 M KCl) and fixative (e.g., 3:1 methanol:glacial acetic acid). Once fixed, cells can be stored almost indefinitely in the fixative solution at 4°C to −20°C.

Slide preparation: Slides are prepared using standard methods, and the cells stained using a suitable staining method that allows good differentiation of sister chromatids (e.g., the fluorescence-plus-Giemsa method of Perry and Thomson, [23]. The slides are then coverslipped to protect the cells, and stored for scoring.

9.4.3.3. Scoring. Uniform criteria for scoring SCEs, such as that suggested by Carrano and Natarajan [19] should be used. Where multiple scorers are involved, a balanced scoring design must be used. Slides must be coded before microscopic analysis and scored in a ''blind'' manner without knowledge of the exposure status of each subject. Slides from the quality assurance standard should be included within this set of slides. Where possible, the number of metaphase cells to be analysed should be based upon the frequency of SCEs usually observed in referents, or on expectations based on historical control data. As this is not always possible, scoring 30–50 second-division metaphases cells per subject is necessary to obtain a stable estimate of the mean to be used in the statistical analysis. The minimal number of scored cells recommended to identify the possible presence of a small proportion (~10%) of high-frequency SCE cells is 80 cells per subject [24]. The latter approach is preferable as it also provides greater confidence in the calculated mean/median value.

The analysis should be carried out by a trained and experienced microscopist who is completely familiar with SCE analysis. Metaphase cells for analysis should be sought under low power magnification and those that are apparently intact with well spread chromosomes should be examined using a high power oil objective (at least 63× magnification). Only second division metaphases, identifiable by their uniform differential staining pattern, containing 46 centromeres are analysed. Every switch of staining between the sister chromatids is scored as an SCE. Switch of label at the centromere can bias the results and should not be scored as an SCE.

In addition to measures of SCEs, information obtained on the relative frequencies of first-division, second-division, and subsequent division metaphase cells, based on BrdU-dependent differential staining patterns, can be used to evaluate cell proliferation kinetics [25,26]. Replication index (RI) [Proliferation index], the average number of replications completed by the metaphase cells examined, can be calculated as follows:

$$\mathrm{RI} = [(\%\,\text{first division metaphases}) + 2 \times (\%\,\text{second division metaphases}) + 3 \times (\%\,\text{subsequent division metaphases}]/100$$

Such information, although not indicative of genotoxic damage, is useful in identifying exposure-induced alterations in mitogen responsiveness and/or the subsequent rate of cell division [27].

9.4.3.4. Data recording. For each second-division metaphase analysed, the number of SCEs, the number of chromosomes, and microscope co-ordinates should be recorded. When a metaphase finder is available for the analysis, the location (and in some models image) of all cells scored and the findings can be stored electronically, facilitating optimal quality assurance. The proportion of first-, second-, and third or subsequent division metaphases is recorded to determine the RI.

9.4.4. Sources of assay variation

The type of culture medium used can affect cell cycle progression, SCE frequency, and replication index.

The type/batch of serum may influence the frequency of SCEs. The type of mitogen influences the proliferation rate of lymphocytes. The stability of the temperature control of incubators, and the temperature variation within the incubators can affect differentially the growth of cultures.

A lack of uniform collection, shipping, processing, and storage of blood samples will increase technical variability.

If cells are stored in fixative, the conditions should be the same for all samples. Variation among scorers is important in all microscopy, including SCE analy-

ses. Variation can be decreased by having only one scorer, and, when several scorers are used, reduced by ensuring a balanced scoring design. Thus, e.g., each scorer should analyse the same number of cells from the slides of all subjects (the same cells not being scored twice) rather than one scorer analysing all cells from some subjects and another scorer all cells from other subjects. Variation can also be reduced by ensuring that the scoring is not unnecessarily protracted across time. Each scorer should be experienced so that scoring criteria do not evolve during the analysis.

9.4.5. Statistical analysis

As described in Section 2, data should be evaluated according to the statistical plan of the study. For each subject, the SCE frequency per cell is calculated. There is no unified agreement on the proper statistical methods to use. However, the unit of exposure is the individual and not the cell; pooling cells across individuals within a study group eliminates a critically important source of variability. The choice of statistical method to use may depend on the distribution of the data. If the distribution of the data is other than normal, methods based on these distributions (e.g., Poisson, binomial, negative binomial, etc.) or non-parametric statistical methods could be used. Alternatively, parametric methods based upon normal distribution can also be used, when appropriate after relevant data transformation. Statistical advice may be needed to determine whether parametric or non-parametric methods are more appropriate for a particular study. The relative and absolute effects of exposure and its confidence interval should be calculated. The confidence interval will complement the negative or positive conclusion by estimating the precision of the effects seen in the study. Stratified analyses can be used for exploring the effect modification by factors such as age, sex, smoking, other exposures, etc. More sophisticated multivariate methods may be necessary for investigating the effect of factors that either may modify or be confounded with the effects [28–30]. Whenever possible, the analysis should be based on individual exposure information. Consideration for the effect of multiple comparisons on the false positive rate should be included in the statistical analysis.

The proportion of HFCs can be used as an additional parameter of genotoxicity [24,30]. Alternatively, the dispersion test (H = variance/mean) can be used to compare the distribution of SCE data among individuals [31].

The quality assurance standard is compared against the appropriate concurrent control cultures using appropriate statistics.

9.4.6. Interpretation of results

Criteria for identifying a positive response (i.e., increased levels of SCEs in an exposed group compared to the referent population; a dose response relationship if variable exposure levels have been established) must be identified prior to implementing the study.

A statistically significantly higher frequency of SCEs among the exposed population in comparison to the referent population is considered a positive result (i.e., the agent of interest induced DNA damage as errors of replication on a damaged DNA template in lymphocytes of exposed individuals). A positive result is supported by a dose response relationship when individuals with different exposure levels are considered. Maximum confidence that the exposure results in an increased frequency of SCEs requires reproducible results in independent studies. It is appreciated that independent studies are not always possible (e.g., in the case of accident situations). Presentation of the results in terms of the size of the effect with confidence intervals may also help with the interpretation.

The lack of a statistically significant increase in SCEs in lymphocytes indicates that under the exposure conditions evaluated and for the calculated power of the study, the exposure did not result in a significant increase in DNA damage in lymphocytes in the population evaluated.

The results of the quality assurance control are used to demonstrate the adequacy of the methodology and the scorers in identifying SCE.

9.5. Report

All data should be presented in tabular or graphic form, and include all observed results (see below). This will allow comparisons between laboratories, re-analyses of results by outside reviewers, and pro-

vide a database for baseline background values and for exposures to physical and chemical agents. All results of the study should be reported with none omitted, and the definition and handling of ''outliers'' described.

The study report should include the following minimal information:

- copy of the questionnaire used to determine personal characteristics
- statement of ethical approval
- characteristics of the exposed and referent populations
- exposure information and analytical methods used
- tissue collection procedures, including timing in relationship to exposure, transportation, storage, and chain-of-custody
- method for coding samples prior to analysis
- assay conditions for cell cultures: number of cultures and replicates, composition of medium, CO2 concentration, incubation duration, colchicine or Colcemid(concentration and duration
- details of the technique used for slide preparation
- number of cells scored per culture and number of replicates
- number of metaphases analysed (data given separately for each subject)
- criteria for scoring SCEs
- criteria for cell kinetic analysis (if conducted)
- number of scorers used and the scoring design
- frequency of SCEs per cell for each subject
- kinetic data for each subject (if scored)
- tabulated results for each subject
- statistical methods
- discussion and interpretation of results
- internal and external quality assurance procedures.

9.6. Literature cited

1. R.R. Tice, A. Hollaender, Sister Chromatid Exchanges: Twenty-five Years of Experimental Research, vol. A, The Nature of Sister Chromatid Exchanges, Plenum, New York, 1984, 491 pp.
2. R.R. Tice, A. Hollaender, Sister Chromatid Exchanges: Twenty-five Years of Experimental Research, vol. B, Genetic Toxicology and Human Studies, Plenum, New York, 1984, 542 pp.
3. J.D. Tucker, A. Auletta, M.C. Cimino, K.L. Dearfield, D. Jacobson-Kram, R.R. Tice, A.V. Carrano, Sister-chromatid exchange: second report of the Gene-Tox program, Mutat. Res. 297 (1993) 101–180.
4. R.B. Painter, A replication model for sister-chromatid exchange, Mutat. Res. 70 (1980) 337–341.
5. J.D. Tucker, R.J. Preston, Chromosome aberrations, micronuclei, sister chromatid exchanges, and cancer risk assessment, Mutat. Res. 365 (1996) 147–159.
6. J.W. Yager, C.J. Hines, R.C. Spear, Exposure to ethylene oxide at work increases sister chromatid exchanges in human peripheral lymphocytes, Science 219 (1983) 1221–1223.
7. J.W. Yager, W.M. Paradisin, S.M. Rappaport, Sister-chromatid exchanges in Lymphocytes are increased in relation to longitudinally measured occupational exposure to low concentrations of styrene, Mutat. Res. 319 (1993) 155–165.
8. L. Hagmar, A. Brogger, I.-L. Hansteen, S. Heim, B. Högstedt, L. Knudsen et al., Cancer risk in humans predicted by increased levels of chromosomal aberrations in lymphocytes: Nordic Study Group on the Health Risk of Chromosome Damage, Cancer Res. 54 (1994) 2919–2922.
9. L. Hagmar, S. Bonassi, U. Stromberg, A. Brogger, L.E. Knudsen, H. Norppa, C. Reuterwall, Chromosomal aberrations in lymphocytes predict human cancer: a report from the European Study Group on Cytogenetic Biomarkers and Health (ESCH), Cancer Res. 58 (1998) 4117–4121.
10. J.D. Tucker, L.K. Ashworth, G.R. Johnston, N.A. Allen, A.V. Carrano, Variation in human lymphocyte sister-chromatid exchange frequency: results of a long-term longitudial study, Mutat. Res. 204 (1988) 435–444.
11. M.A. Bender, R.J. Preston, R.C. Leonard, B.E. Pyatt, P.C. Gooch, Chromosomal aberrations and sister-chromatid exchanges in peripheral lymphocytes of a large human population sample. II Extension of age range, Mutat. Res. 212 (1989) 149–154.
12. J.R. Lazutka, V. Dedonyte, D. Krapavickaite, Sister-chromatid exchanges and their distribution

in human lymphocytes in relation to age, sex and smoking, Mutat. Res. 306 (1994) 173–180.
13. C. Bolognesi, A. Abbondandolo, R. Barale, R. Casalone, L. Dalpra, M. De Ferrari et al., Age-related increase of baseline frequencies of sister chromatid exchanges, chromosome aberrations, and micronuclei in human lymphocytes, Cancer Epidemiol. Biomarkers Prev. 6 (1997) 249–256.
14. R. Barale, L. Chelotti, T. Davini, S. Del Ry, M.G. Andreassi, M. Ballardin et al., Sister chromatid exchange and micronucleus frequency in human lymphocytes of 1,650 subjects in an Italian population: II. Contribution of sex, age, and lifestyle, Environ. Mol. Mutagen. 31 (1998) 228–242.
15. J.K. Wiencke, S. Pemble, B. Ketterer, K.T. Kelsey, Gene deletion of glutathione *S*-transferase theta: correlation with induced genetic damage and potential role in endogenous mutagenesis, Cancer Epidemiol. Biomarkers Prev. 4 (1995) 254–259.
16. K.T. Kelsey, J.K. Wiencke, J. Ward, W. Bechtold, J. Fajen, Sister chromatid exchanges, glutathione *S*-transferase theta deletion and cytogenetic sensitivity to diepoxybutane in lymphocytes from butadiene monomer production workers, Mutat. Res. 335 (1995) 267–273.
17. H. Norppa, A. Hirvonen, H. Järventaus, M. Uuskylä, G. Tasa, A. Ojajärvi, M. Sorsa, Role of GSTT1 and GSTM1 genotypes in determining individual sensitivity to sister chromatid exchange induction by diepoxybutane in cultured human lymphocytes, Carcinogenesis 16 (1995) 2116–2119.
18. K.R. Schröder, F.A. Wiebel, P. Danappel, H.M. Bolt, E. Hallier, Gluthione *S*-transferase (GST) theta polymorphism influences background SCE rate, Arch. Toxicol. 69 (1995) 505–507.
19. A.V. Carrano, A.T. Natarajan, Considerations for population monitoring using cytogenetic techniques, International Commission for protection against environmental mutagens and carcinogens. ICPEMC Publication No. 14, Mutat. Res. 204 (1988) 379–406.
20. WHO, Safety in Health-Care Laboratories, World Health Organization, Geneva, 1997, 148 pp.
21. US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, and National Institutes of Health. Biosafety in Microbiological and Biomedical Laboratories. HHS Publication No (CDC) 93-8395. 3rd edn. (May 1993), 169 pp., as printed as an Appendix of D.O. Fleming, J.H. Richardson, J.J. Tulis, D. Vesley (Eds.), Laboratory Safety, Principles and Practices, 2nd edn., ASM Press, Washington DC, 1995.
22. P. Perry, L. Henderson, D. Kirkland, Sister Chromatid Exchange in Cultured Cells, Report of the UKEMS Subcommittee on Guidelines for Mutagenicity Testing, Part 11, 1984, pp. 89–122.
23. P.E. Perry, E.J. Thomson, The Methodology of Sister Chromatid Exchange, in: B.J. Kilbey, M. Legator, W. Nichols, C. Ramel (Eds.), Handbook of Mutagenicity Test Procedures, 2nd edn., Elsevier Scientific, Amsterdam, 1984, pp. 495–530.
24. D.H. Moore, II, A.V. Carrano, Statistical analysis of high SCE frequency cells in human lymphocytes, in: R.R. Tice, A. Hollaender (Eds.), Sister Chromatid Exchanges: Twenty-five Years of Experimental Research. vol. A. The Nature of Sister Chromatid Exchanges, Plenum, New York, NY, 1984, pp. 469–479.
25. R. Tice, E.L. Schneider, J.M. Rary, The utilization of bromodeoxyuridine incorporation into DNA for the analysis of cellular kinetics, Exp. Cell Res. 102 (1976) 232–236.
26. R.R. Tice, P. Thorne, E.L. Schneider, BISACK analysis of the phytohemagglutinin-induced proliferation of human peripheral lymphocytes, Cell Tissue Kinet. 12 (1979) 1–9.
27. V. Palma, H. Tudon, L. Buentello, S. Nava, P. Ostrosky, F. Salamanca, Methods for the analysis of cellular kinetics in PHA-stimulated blood lymphocytes using BrdU incorporation. A comparative study, Mutat. Res. 286 (1993) 267–273.
28. C. Bolognesi, A. Abbondandolo, R. Barale, R. Casalone, L. Dalpra, M. De Ferrari et al., Age-related increase of baseline frequencies of sister chromatid exchanges, chromosome aberrations, and micronuclei in human lymphocytes, Cancer Epidemiol. Biomarkers Prev. 6 (1997) 249–256.
29. S. Bonassi, M. Ceppi, V. Fontana, F. Merlo, Multiple regression analysis of cytogenetic human data, Mutat. Res. 313 (1994) 69–80.

30. S. Bonassi, V. Fontana, M. Ceppi, R. Barale, A. Biggeri, Analysis of correlated data in human biomonitoring studies. The case of high sister chromatid exchange frequency cells, Mutat. Res. 438 (1999) 13–21.

31. B. Margolin, M.D. Shelby, Sister chromatid exchanges: A re-examination of the evidence for sex and race differences in human, Environ. Mutagen. 7 (1985) 63–72.