An ECVAG inter-laboratory validation study of the comet assay: inter-laboratory and intra-laboratory variations of DNA strand breaks and FPG-sensitive sites in human mononuclear cells.

Ersson C;Møller P;Forchhammer L;Loft S;Azqueta A;Godschalk RW;van Schooten FJ;Jones GD;Higgins JA;Cooke MS;Mistry V;Karbaschi M;Phillips DH;Sozeri O;Routledge MN;Nelson-Smith K;Riso P;Porrini M;Matullo G;Allione A;Stepnik M;Ferlinska M;Teixeira JP;Costa S;Corcuera LA;López de Cerain A;Laffon B;Valdiglesias V;Collins AR;Möller L. An ECVAG inter-laboratory validation study of the comet assay: inter-laboratory and intra-laboratory variations of DNA strand breaks and FPG-sensitive sites in human mononuclear cells.. MUTAGENESIS. 28 (3) pp: 279-286.
DOI: 10.1093/mutage/get001

An ECVAG inter-laboratory validation study of the comet assay: inter-laboratory and intra-laboratory variations of DNA strand breaks and FPG-sensitive sites in human mononuclear cells

Clara Ersson, Peter Møller, Lykke Forchhammer, Steffen Loft, Amaya Azqueta Roger W. L. odschalk, Frederik-Jan van Schooten, George D. D. Jones, Jennifer A. Higgins, Marcus S. Cooke, Vilas Mistry, Mahsa Karbaschi, David H. Phillips, Osman Sozeri, Michael N. Routledge, Kirsty Nelson-Smith, Patrizia iso, Marisa Porrini, Giuseppe Matullo, Alessandra Allione, Maciej Stępnik, Magdalena Verlińska, João Paulo Teixeira, Solange Costa, Laura-Ana Corcuera, Adela López de Cerain, Blanca Laffon, Vanessa Valdiglesias, Andrew R. Collins, Lennart Möller

Abstract

The alkaline comet assay is an established, sensitive method extensively used in biomonitoring studies. This method can be modified to measure a range of different types of DNA damage. However, considerable differences in the protocols used by different research groups affect the inter-laboratory comparisons of results. The aim of this study was to assess the inter-laboratory, intra-laboratory, sample and residual (unexplained) variations in DNA strand breaks and formamidopyrimidine DNA glycosylase (FPG)-sensitive sites measured by the comet assay by using a balanced Latin square design. Fourteen participating laboratories used their own comet assay protocols to measure the level of DNA strand breaks and FPG-sensitive sites in coded samples containing peripheral blood mononuclear cells (PBMC) and the level of DNA strand breaks in coded calibration curve samples (cells exposed to different doses of ionising radiation) on three different days of analysis. Eleven laboratories found dose–response relationships in the coded calibration curve samples on two or three days of analysis, whereas three laboratories had technical problems in their assay. In the coded calibration curve samples, the dose of ionising radiation, inter-laboratory variation, intra-laboratory variation and residual variation contributed to 60.9, 19.4, 0.1 and 19.5%, respectively, of the total variation. In the coded PBMC samples, the inter-laboratory variation explained the largest fraction of the overall variation of DNA strand breaks (79.2%) and the residual variation (19.9%) was much larger than the intra-laboratory (0.3%) and inter-subject (0.5%) variation. The same partitioning of the overall variation of FPG-sensitive sites in the PBMC samples indicated that the inter-laboratory variation was the strongest contributor (56.7%), whereas the residual (42.9%), intra-laboratory (0.2%) and inter-subject (0.3%) variations again contributed less to the overall variation. The results suggest that the variation in DNA damage, measured by comet assay, in PBMC from healthy subjects is assay variation rather than variation between subjects.

Introduction

The alkaline comet assay is an established, sensitive method extensively used in biomonitoring studies. In the method's basic form, DNA damage (including DNA strand breaks and alkali-labile sites) is measured. This method can also be modified to measure a range of specific types of DNA damage by adding lesion-specific repair enzymes (1). For example, oxidatively damaged DNA can be measured by adding formamidopyrimidine DNA glycosylase (FPG), which recognises and removes oxidatively damaged purines, for example, 8-oxo-7,8-dihydroguanine. Other modifications of the comet assay allow the measurement of DNA–DNA cross-links (2), detection of DNA damage in specific genes (3) and activity of DNA repair (4–7). The many applications of the comet assay are important reasons for the assay's popularity. Other advantages of the method include the low number of cells required (as little as a few thousand), the ability to measure heterogeneity in sensitivity or response between cells, the sensitivity in detecting low levels of DNA damage and the potential to measure damage in practically any eukaryotic cell type, including tissues (8).

Several guidelines for the comet assay have been published (9–13). However, as yet, there is no generally accepted standard protocol for the comet assay and there are considerable differences in the protocols used by different research groups. Because the protocol used may substantially affect the outcome of the

comet assay, this introduces serious problems when attempting inter-laboratory comparisons of results (14–17). Although this is relatively well known, few studies have addressed the issue. Several experts of the comet assay have emphasised that before comet assay biomarkers can be considered as useful tools in cancer risk assessment, inter-laboratory and intra-laboratory reproducibility and reliability, and the sources of variability should be assessed in multi-laboratory, international validation studies (11, 13, 18). The European Comet assay Validation Group (ECVAG) has previously demonstrated that the protocols used affect the outcome of the comet assay and that this variation can be decreased by using calibration curve samples, internal standards or a standardised protocol (16, 17, 19–21). By analysing samples of cultured cells that had been treated with ionising radiation, we demonstrated that the intra-laboratory variation was much lower than the inter-laboratory variation (16). However, that analysis did not discriminate between the within-day and between-day variation in the individual laboratories. In addition, the use of ionising radiation in cultured cells might not be relevant for the assessment of the variation of DNA damage in blood cells from healthy donors.

The aim of this study was to assess the inter-laboratory, intra-laboratory, sample and unexplained (residual) variations of DNA strand breaks and FPG-sensitive sites in peripheral blood mononuclear cells (PBMC) measured by the comet assay by using a balanced, so-called Latin square design. The participating laboratories analysed the same three differently coded samples of PBMC from each of three different donors and coded calibration curve samples, and the analysis was repeated on three separate days of analysis.

Materials and methods

Study design

Fourteen laboratories participated in this inter-laboratory validation study. Each participating laboratory received FPG enzyme, 27 coded samples of PBMC from three different female donors (aged 24–45 years, nine samples from each donor) and three sets of coded calibration curve samples that were to be analysed by their own comet assay procedure on three different days, that is, three aliquots from each donor were to be analysed on each day of analysis. The calibration curve samples were analysed in the same experiments (electrophoresis runs) as the donor samples, and therefore, the calibration curve samples can also be considered to be 'reference standards'. Each aliquot had a different code, and it was not possible to identify which samples came from the same donor. By using a Latin square design (analysis of three donors and three identical aliquots per donor on three different days), discrimination between inter-laboratory, intra-laboratory, sample and residual variations could be performed. Ethical permission for the study was received from the ethical committee in Stockholm, Sweden (2009/629–31/4 and 2009/974-32).

Reagents and enzymes

Roswell Park Memorial Institute 1640 medium and foetal bovine serum were purchased from Invitrogen Corporation (Karlsruhe, Germany). FPG was supplied by one of the participating laboratories (Department of Nutrition, University of Oslo, Norway).

Preparation of coded calibration curve samples

Calibration curve samples, exposed to 0, 2.5, 5 or 7.5 Gy of γ-rays, were prepared and frozen as described by Forchhammer *et al.* (19).

Preparation of coded PBMC samples

Venous blood samples were collected in BD Vacutainer® CPT™ cell preparation tubes containing sodium heparin anticoagulant (Becton-Dickinson, Franklin Lakes, NJ, USA), which allow direct isolation of

mononuclear cells from whole blood by centrifugation. Aliquots of PBMC were frozen slowly in freezing media consisting of 50% foetal bovine serum, 40% Roswell Park Memorial Institute 1640 culture medium and 10% dimethyl sulphoxide (DMSO).

The comet assay

The 14 participating laboratories used their own comet assay protocols to measure the level of DNA strand breaks and FPG-sensitive sites in the coded samples. Laboratory-specific conditions are specified in Table I. The basic procedure was common for all laboratories. Cells were embedded in 0.49–1% low-melting-point agarose (in phosphate-buffered saline) and then lysed for a minimum of 1h at 4°C. The lysis buffer generally consisted of 1% Triton X-100, 2.5M NaCl, 0.1M EDTA, 10mM Tris, pH 10. In addition, the lysis solution of Laboratory 8 contained 1% DMSO and the lysis solution of Laboratory 14 contained 250mM NaOH (pH > 12). Laboratory 11 used a different lysis buffer (1% Triton X-100, 2.5mM NaCl, 1mM EDTA, 10mM Tris, 10% DMSO, pH 10). After lysis, the gels were washed in buffer (40mM HEPES, 0.1M KCl, 0.5mM EDTA, 0.2mg/ml bovine serum albumin, pH 8) and then incubated with only buffer or with buffer containing FPG. Incubation was performed at 37°C for 30–45min, depending on the laboratory. This enabled an estimation of both DNA strand breaks and net FPG-sensitive sites. In the case of the calibration curve samples, incubation was omitted and slides were transferred directly from lysis to an alkaline solution (0.3M NaOH, 1mM EDTA, pH > 13). Slides were placed in the alkaline solution for 20–40min, followed by electrophoresis in the same solution for 20–30min. The DNA was stained and the comets were analysed either by visual scoring or by computerised image analysis.

Table I. Own comet assay protocols used by the different laboratories

Lab	Agarose LMP in PBS (%)	Enzyme incubation (min)	Alkali (min)	Electrophoresis				Stain[b]	Software
				Time (min)	Voltage (V)	V/cm[a]	Amperage (mA)		
1	0.6	30	20	20	30	1.2	300	EtBr	Comet IV (Perceptive Instruments)
2	1	30	20	30	35	1.1	No info	DAPI	Comet IV (Perceptive Instruments)
3	0.8	30	20	20	25	1.7	285–300	SYBR® Gold	Comet IV (Perceptive Instruments)
4	0.75	45	40	20	30	0.8	300	YOYO-1	Comet IV (Perceptive Instruments)
5	0.75	45	40	20	25	1.1	300	YOYO-1	-
6	0.7	30	30	30	25	0.7	300	DAPI	Comet IV (Perceptive Instruments)
7	0.49	30	20	24	22	1.4	300	EtBr	Comet IV (Perceptive Instruments)
8	1	45	40	20	25	1.1	300	EtBr	Cometa 1.5 (Immagini e Computer, Italy)
9	0.6	30	20	20	27	0.9	300	PI	Comet IV (Perceptive Instruments)
10	0.6	30	20	20	30	0.9	300	PI	Komet 5.5 (Kinetic Imaging Ltd)
11	0.8	30	40	20	23	0.9	300	EtBr	Komet 5.5 (Kinetic Imaging Ltd)
12	1	30	20	30	25	0.9	300	DAPI	LUCIA 4.61 (Precoptic C, Czech Rebublic)
13	0.7	30	40	30	17	1.2	300	EtBr	Komet 4.0 (Kinetic Imaging Ltd)
14	0.65	30	20	20	25	1.1	300	EtBr	Comet III (Perceptive Instruments)

LMP, low melting point; PBS, phosphate-buffered saline. [a]cm, length of electrophoresis platform. [b]The dyes are ethidium bromide (EtBr), 4′6, diamidino-2-phenylindole dihydrochloride (DAPI), propidium iodide (PI) and YOYO®-1 iodide (YOYO-1).

Reported endpoints

We have reported primary comet assay endpoints obtained by computerised scoring as %DNA in the tail (*n* = 13), and visual scoring as arbitrary units (a.u.) in the scale 0–100 (*n* = 1). Visual scoring and computerised scoring have been shown to correlate to each other (18, 22, 23). These primary comet assay endpoints can subsequently be transformed to lesions/10^6 bp by using calibration curves. Based on the assumptions that 1 Gy of γ-rays induces 0.29 strand breaks/10^9 Da DNA [the mean of the levels reported by Ahnström *et al.* (24) and Kohn *et al.* (25)] and that 1bp has a mass of 650Da, the slopes of the calibration curves (%DNA in tail/Gy) were used to calculate how many lesions/10^6 bp a particular percentage of DNA in the tail corresponded to

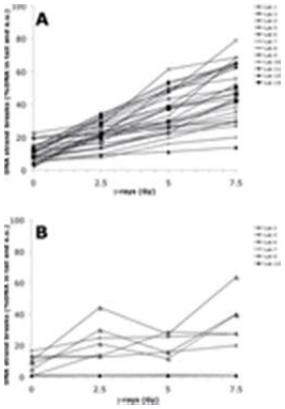
$$\text{Number of lesions}/10^6\ \text{bp} = \frac{\%\text{DNA in tail}_{\text{sample}}}{\text{Slope of calibration curve}} \times \frac{0.29}{10^9} \times 650 \times 10^6.$$

The slope of the calibration curve refers to the slope of each specific calibration curve.

Statistics

We have analysed the results in different strata of the dataset according to the completeness of the dose–response relationships of the calibration curve samples. Three laboratories obtained non-significant or no dose–response relationship in the calibration curve samples on two or three days of analysis (Laboratories 2, 12 and 14), and the data from these laboratories were excluded from all subsequent statistical analyses. These exclusion criteria also included the conversion to lesions/10^6 bp because the slopes of the calibration curves are used in the conversion. The non-significant calibration curves for Laboratories 2 and 12 are presented in Figure 1B. Laboratory 14 did not report any calibration curves due to the problem of there being too few nucleoids present in the gels.

Figure 1.



The mean levels of DNA strand breaks in cells exposed to 0, 2.5, 5 and 7.5 Gy. Investigators from 14 laboratories have measured the level of %DNA in tail (*n* = 13) and a.u. (scale 0–100, *n* = 1) on three different days of analysis using their own comet assay procedure. Laboratory 14 did not report any data for the calibration curves. Data have been divided into (**A**) analyses where significant linear dose–response relationships were obtained (R^2_{mean} = 0.968, range: 0.908–0.999, *n* = 30, *P* < 0.05) and (**B**) analyses where the investigators failed to detect a significant linear dose–response relationship (linear regression). Note that for each laboratory the dose–response relationships have been analysed separately for each of the three days of analysis.

Statistical analyses were performed both with and without results from four laboratories that did succeed in detecting a significant dose–response relationship on two days of analysis but which failed to detect a significant dose–response relationship on one day of analysis (Laboratories 4, 6, 7 and 9). One stratum

contained the results from the 11 laboratories that had obtained two or three full dose–response curves on two or all three days (Laboratories 1, 3–11 and 13). Note that in Figure 2 and Table IV, all three calibration curves were used to convert the primary comet assay endpoints to lesions/10^6 bp for these 11 laboratories (including the non-significant calibration curves of Laboratories 4, 6, 7 and 9). The slopes of these laboratories' non-significant and significant calibration curves were similar. Laboratory 7 failed to detect the dose–response relationship due to lost gels and Laboratories 4, 6 and 9 reported dose–response relationships for the calibration curve samples that were not far from being significant (P = 0.05–0.13). In a second stratum, only the seven laboratories that obtained significant dose–response relationships on all three days of analyses were included (Laboratories 1, 3, 5, 8, 10, 11 and 13). Note that for each laboratory, the dose–response relationships have been analysed separately for each of the three days of analysis.

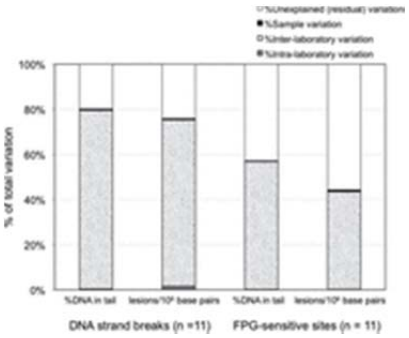
Table IV. Mean levels of DNA strand breaks and FPG-sensitive sites (presented as %DNA in tail and lesions/10^6 bp) in the coded PBMC samples and CVs obtained in each laboratory

Lab	Person	DNA strand breaks				FPG-sensitive sites			
		%DNA in tail	CV	Lesions/10^6 bp	CV	%DNA in tail	CV	Lesions/10^6 bp	CV
1	1	24.87	0.29	1.17	0.29	11.69	0.36	0.54	0.32
	2	20.21	0.45	0.95	0.43	11.77	0.31	0.57	0.38
	3	19.14	0.38	0.93	0.43	6.48	0.55	0.30	0.49
2	1	6.80	1.98	NA	NA	26.05	0.87	NA	NA
	2	4.35	1.12	NA	NA	18.31	0.92	NA	NA
	3	9.51	1.37	NA	NA	25.05	0.80	NA	NA
3	1	1.76	0.35	0.06	0.32	7.90	0.54	0.29	0.52
	2	1.11	0.30	0.04	0.32	7.57	0.52	0.28	0.53
	3	1.86	0.60	0.07	0.58	8.46	0.37	0.31	0.37
4	1	6.41	0.34	0.63	0.46	9.20	0.39	0.89	0.45
	2	4.61	0.22	0.44	0.30	7.69	0.33	0.73	0.39
	3	4.14	0.33	0.40	0.44	8.28	0.21	0.79	0.28
5	1	2.83	0.63	0.06	0.73	13.83	0.44	0.31	0.54
	2	1.22	0.97	0.03	1.08	13.18	0.79	0.28	0.74
	3	2.26	0.51	0.05	0.57	11.50	0.51	0.25	0.51
6	1	10.34	0.17	1.11	0.50	7.68	0.50	0.70	0.49
	2	9.51	0.33	0.96	0.47	7.78	0.55	0.72	0.66
	3	9.82	0.17	1.02	0.45	5.66	0.57	0.55	0.60
7	1	0.88	0.33	0.02	0.44	10.42	0.84	0.29	1.06
	2	0.67	0.50	0.02	0.56	12.00	0.85	0.35	1.06
	3	0.70	0.53	0.02	0.67	8.87	0.81	0.25	1.01
8	1	6.68	0.15	0.18	0.14	23.98	0.12	0.64	0.10
	2	7.28	0.18	0.19	0.17	23.67	0.11	0.63	0.12
	3	7.24	0.18	0.19	0.16	22.72	0.13	0.61	0.13
9	1	7.58	0.33	0.64	0.21	5.17	0.57	0.45	0.51
	2	9.16	0.37	0.75	0.22	4.11	0.78	0.36	0.73
	3	5.42	0.36	0.47	0.11	3.33	1.12	0.24	1.01
10	1	7.10	0.22	0.50	0.23	1.71	1.42	0.12	1.28
	2	7.77	0.15	0.56	0.25	2.99	0.61	0.20	0.52
	3	7.36	0.14	0.53	0.30	3.12	0.80	0.20	0.68
11	1	13.24	0.18	0.36	0.18	15.43	0.62	0.42	0.58
	2	11.00	0.07	0.30	0.06	13.77	0.51	0.37	0.48
	3	12.97	0.16	0.36	0.14	18.47	0.65	0.50	0.62
12	1	0.97	0.41	NA	NA	1.97	0.94	NA	NA
	2	1.51	0.67	NA	NA	2.05	1.07	NA	NA

Lab	Person	DNA strand breaks				FPG-sensitive sites			
		%DNA in tail	CV	Lesions/10^6 bp	CV	%DNA in tail	CV	Lesions/10^6 bp	CV
	3	1.69	0.37	NA	NA	0.92	1.62	NA	NA
	1	6.70	0.29	0.26	0.27	18.18	0.10	0.70	0.10
13	2	5.85	0.09	0.23	0.10	18.72	0.14	0.72	0.10
	3	6.60	0.18	0.26	0.18	18.18	0.10	0.71	0.15
	1	14.37	0.48	NA	NA	26.43	0.48	NA	NA
14	2	14.48	0.32	NA	NA	30.63	0.53	NA	NA
	3	13.20	0.40	NA	NA	26.85	0.44	NA	NA

The coded samples of PBMC were obtained from three different subjects. Each laboratory analysed three coded samples from each subject at three different days, that is, each mean value is based on the analyses of nine coded samples. NA means non-applicable and refers to that the lesions/10^6 bp cannot be calculated without calibration curves.

Figure 2.



The percentage of the overall variation explained by the different PBMC samples, the inter-laboratory variation, the intra-laboratory variation and the unexplained (residual) variation of the reported levels of DNA strand breaks and FPG-sensitive sites measured by the comet assay (three-way ANOVA). The experiment had a Latin square design, that is, the participating laboratories analysed three identical, but differently coded, samples of PBMC from each of three different donors and calibration curve samples on three different days. The statistics is based on data from the 11 laboratories that reported significant dose–response relationships in the calibration curve samples on two or three days of analysis.

Statistical analyses were performed using PASW Statistics 18.0 (SPSS Inc., Chicago, IL, USA). The dose–response relationships of the calibration curves were analysed by linear regression analyses. The part of the variation attributed to the day of analysis, the laboratory and the dose or samples was analysed by analysis of variance (ANOVA) with the day of analysis as a categorical random variable, laboratory as a categorical fixed factor and the PBMC sample or dose of γ-radiation as fixed or continuous variables. The ANOVA model did not test for interactions between variables. The assessment of the contribution of each variable to the overall total variation was calculated from the sum of squares in the statistical analyses, and the fraction of the variation that was not explained by these factors was the residual variation. The residual variation is the unexplained part of the total variation in the dataset.

Results

Calibration curve samples

The mean levels of DNA strand breaks in the calibration curve samples measured as either %DNA in tail (n = 12) or a.u. (scale 0–100, n = 1) are shown in Figure 1. Figure 1A depicts the analyses where significant linear dose–response relationships were obtained (R^2_{mean} = 0.968, range: 0.908–0.999, n = 30, $P < 0.05$) and Figure 1B depicts the analyses where the investigators failed to detect a significant linear dose–response relationship (linear regression). Seven of the participating laboratories (Laboratories 1, 3, 5, 8, 10, 11 and 13) reported significant dose–response relationships on all three days of analysis. Four laboratories (Laboratories 4, 6, 7 and 9) reported significant dose–response relationships on two days of analysis and

failed to detect the dose–response relationship on one occasion. Of these laboratories, Laboratory 7 failed to detect the dose–response relationship due to lost gels and Laboratories 4, 6 and 9 reported non-significant dose–response relationships for the calibration curve samples (P = 0.05–0.13). Laboratory 12 failed to detect a dose–response relationship on two days of analysis and Laboratory 2 failed to detect the dose–response relationship on all three days of analysis. Laboratory 14 did not report any data for the calibration curves. The statistical power to obtain a significant dose–response relationship for the calibration curve samples from a single day of analysis is relatively low. If the dose–response relationships of the calibration curves from all three days of analysis were analysed together instead as for each day separately, all laboratories that reported data for the calibration curves (Laboratories 1–13) could detect a significant dose–response relationship (Table II). In addition, the slopes and the R_{mean} values of the calibration curves analysed on the three different days of analysis are reported in this table. The dose of γ-radiation, inter-laboratory variation, intra-laboratory variation and unexplained (residual) variation contributed 60.9 (P < 0.001), 19.4 (P < 0.001), 0.1 (P > 0.05) and 19.5%, respectively, of the total variation. Although the intra-laboratory variation was low in such an overall analysis, it can be seen in Table II that the intra-laboratory variation was significant in four of the participating laboratories.

Table II. Regression coefficient (R) and slope of calibration curves from three different days of analysis and statistical significance of the dose–response relationship

ab	Day	*R* value	Slope	Intercept	Effect (*P* value)[a]		
					Dose	Day	Dose*Day
	1[b]	0.97	4.9	5.6			
1	2[b]	0.97	3.5	17.7	<0.001	0.028	0.398
	3[b]	0.98	3.8	21.9			
	1	0.81	5.7	14.8			
2	2	0.76	3.7	8.8	0.023	0.896	0.755
	3	0.77	3.3	7.9			
	1[b]	0.98	5.2	8.4			
3	2[b]	0.95	4.7	4.3	<0.001	0.728	0.900
	3[b]	0.97	5.3	8.8			
	1[b]	0.97	1.7	18.3			
4	2[b]	0.99	2.1	14.7	<0.001	0.330	0.772
	3	0.87	2.3	12.1			
	1[b]	0.97	8.7	8.5			
5	2[b]	0.97	7.4	4.7	<0.001	0.663	0.406
	3[b]	0.99	9.9	1.0			
	1[b]	0.99	3.8	6.5			
6	2[b]	1.00	2.0	11.8	<0.001	0.036	<0.001
	3	0.95	1.1	11.2			
	1[b]	0.99	8.0	2.8			
7	2	1.00	5.6	0.7	0.004	0.888	0.730
	3[b]	0.98	8.4	-0.6			
	1[b]	1.00	7.2	12.0			
8	2[b]	1.00	7.1	12.8	<0.001	0.822	0.768
	3[b]	1.00	6.9	11.5			
	1[b]	0.98	3.2	15.0			
9	2[b]	0.98	2.4	19.4	<0.001	0.312	0.054
	3	0.90	1.3	18.6			
	1[b]	0.97	2.0	4.8			
10	2[b]	0.97	3.2	4.5	<0.001	0.624	0.161
	3[b]	0.99	3.0	7.0			
	1[b]	1.00	6.7	15.3			
11	2[b]	0.99	6.7	15.7	<0.001	0.799	0.755
	3[b]	0.99	7.2	13.3			

ab	Day	*R* value	Slope	Intercept	Effect (*P* value)[a]		
					Dose	Day	Dose*Day
	1	0.82	0.0	0.8			
12	2	0.04	0.0	1.0	<0.001	<0.001	<0.001
	3[b]	0.98	1.2	5.2			
	1[b]	1.00	5.0	13.6			
13	2[b]	0.99	5.2	8.8	<0.001	0.046	0.265
	3[b]	1.00	4.5	8.0			

Laboratory 14 did not report any calibration curves due to problems with too few nucleoids in the gels. [a]The statistical significance of the contribution of the dose of γ-radiation, the day of analysis and the interaction term (dose*day, which tells us whether or not the slopes are different). The statistical significances were obtained from a general linear model analysis, where the %DNA in tail was the dependent variable, the day of analysis was a random categorical factor and the dose of γ-radiation was a continuous variable. [b]A significant dose–response relationship ($P < 0.05$) was observed when the calibration curves obtained at each of the three days of analysis were analysed separately in a simple regression analysis.

Table III depicts the mean values for the calibration curve samples and the standard deviation obtained in each laboratory (note that three laboratories have been removed on the basis of the exclusion criteria). In addition, the percentage of the total variation attributed to the dose of γ-radiation, the intra-laboratory variation and the unexplained (residual) variation, as well as the standard deviation of the residuals, are presented (Table III).

Table III. Assessment of the contribution of intra-laboratory (day-to-day) variation to the overall variation of the three calibration curves[a]

Lab	DNA strand breaks (%DNA in tail and a.u.) reported as mean (SD)								% of total variation[b]			
	0 Gy		2.5 Gy		5 Gy		7.5 Gy		Dose	Intra-laboratory variation	Residual variation	SD_{res}[c]
1	15.0	(10.5)	24.9	(3.7)	36.0	(8.7)	45.0	(3.6)	78.3***	15.5**	6.2	3.3
3	9.9	(1.7)	17.0	(3.8)	30.2	(7.7)	47.8	(4.1)	90.0***	3.9	6.1	3.9
4	15.8	(3.2)	18.1	(4.3)	26.9	(1.7)	29.7	(2.5)	79.1***	6.7	14.2	2.5
5	5.2	(2.2)	23.5	(3.1)	52.7	(7.6)	67.7	(11.6)	93.2***	2.0	4.9	5.8
6	10.8	(2.5)	14.1	(1.8)	21.1	(4.8)	27.4	(7.7)	71.3***	9.6	19.1	3.4
7	1.7	(0.9)	20.3	(5.5)	36.0	(4.3)[d]	66.4	(2.6)[d]	95.7***	0.7	3.6	4.2
8	11.0	(0.7)	30.7	(1.0)	48.9	(1.2)	63.7	(1.7)	99.4***	0.2	0.4	1.4
9	17.2	(2.1)	25.0	(1.2)	27.6	(1.9)	35.6	(7.1)	78.5***	7.4	14.0	2.9
10	6.7	(0.6)	10.0	(2.9)	19.6	(3.7)	26.3	(5.6)	84.5***	8.8*	6.7	2.2
11	13.0	(1.2)	33.3	(1.1)	51.4	(3.0)	64.3	(1.3)	98.5***	0.0	1.5	2.5
13	9.6	(2.4)	22.8	(4.5)	35.2	(4.8)	46.2	(4.2)	94.3***	4.8***	1.0	1.4

[a]The results from Laboratories 2, 12 and 14 are not included in this table based on the exclusion criteria. [b]Analysed by ANOVA in dataset strata from individual laboratories. The ANOVA model included the day of analysis as random factor and the dose of γ-radiation as continuous variable. The data are reported as the percentage of the total sum of squares. [c]SD_{res} is the standard deviation of the residuals. [d]Based on two analysis occasions instead of three. *$P < 0.001$, **$P < 0.01$ and ***$P < 0.05$.

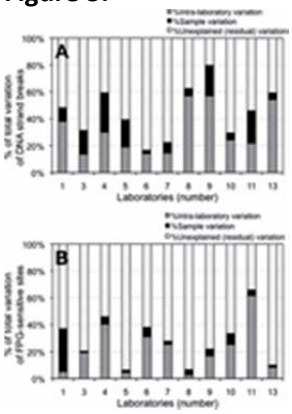
DNA damage in coded samples

The mean levels of DNA strand breaks and FPG-sensitive sites of the coded PBMC samples are presented in Table IV together with the coefficients of variation (CVs) for all participating laboratories. Based on the values obtained from the seven laboratories that obtained significant dose–response relationships for the calibration curve samples on all three days of analysis, the mean levels of DNA strand breaks/10^6 bp in the PBMC samples from the three donors were 0.37 (range: 0.06–1.17; Sample 1), 0.33 (range: 0.03–0.95; Sample 2) and 0.34 (range: 0.05–0.93; Sample 3), and the mean levels of FPG-sensitive sites/10^6 bp were 0.43 (range: 0.12–0.70; Sample 1), 0.44 (range: 0.20–0.72; Sample 2) and 0.41 (range: 0.20–0.71; Sample 3).

The inter-laboratory variation explained the largest fraction of the overall variation in the PBMC and the unexplained (residual) variation was much larger than the intra-laboratory variation (Figure 2). In addition, it can be seen that the inter-laboratory variation is reduced by conversion of the primary comet assay endpoints to lesions/10^6 bp for both the DNA strand breaks and the FPG-sensitive sites.

The contribution of the inter-laboratory variation, the PBMC sample and the unexplained (residual) variation to the overall variation of DNA strand breaks and FPG-sensitive sites (reported in %DNA in tail or a.u.) in each individual laboratory are shown in Figure 3A and B, respectively.

Figure 3.



The percentage of the overall variation explained by the different samples, intra-laboratory variation and unexplained (residual) variation in PBMC samples analysed in 11 laboratories (two-way ANOVA). The data are based on the level of (**A**) DNA strand breaks and (**B**) FPG-sensitive sites, in PBMC reported as %DNA in tail or a.u. in the scale 0–100 with the comet assay.

Discussion

The balanced Latin square design used in this study, whereby the participating laboratories analysed coded samples containing PBMC from three donors and three identical samples per donor on each of three different days, allowed us to discriminate between inter-laboratory, intra-laboratory, sample and residual variations. An important finding in this study is that the inter-laboratory variation explained the largest fraction of the overall variation and that the unexplained (residual) variation was much larger than the intra-laboratory variation (Figure 2). The inter-laboratory variability is likely to be explained by the different comet assay procedures and equipment used by the different research groups. Protocol steps that seem to be of particular importance are the agarose density, the duration of enzyme treatment, the extent of alkaline treatment, the duration of electrophoresis and the strength of the electric field (14, 15, 17). In addition, the findings of one of our previous ECVAG inter-laboratory validation studies implied that the image analysis (and possibly staining) affects the outcome of the method (16). The intra-laboratory variability corresponds to variation in procedures on different days of analysis, whereas the unexplained (residual) variation corresponds to the variability between the three putatively identical samples that were analysed on the same day. Sources of this latter type of variability includes all procedures of the handling of samples, including the original pipetting of aliquots in the laboratory that prepared the samples, handling of samples in the laboratories that analysed the samples (e.g. samples are embedded in the gels one at a time and small differences in the processing would be registered as residual variation) and gel-to-gel differences related to all procedures from lysis to staining/scoring of the nuclei in the gel. A part of the gel-to-gel differences might be attributed to heterogeneity in the strength of the electric field during electrophoresis. To what extent the variation in strength of the electric field contributes to the gel-to-gel or within-gel variation in the comet assay needs to be investigated. It has been observed that the real electric potential (V/cm) as measured using closely spaced electrodes on the platform may differ substantially at various positions and may also vary with time, relative to what can be calculated from the applied electrode voltage (G. Brunborg, personal communication). The mean levels of DNA damage reported in this study (0.35 DNA strand breaks/10^6 bp and 0.43 FPG-sensitive sites/10^6 bp) are similar to the levels we have

previously observed in young healthy women (0.29 DNA strand breaks/10^6 bp and 0.48 FPG-sensitive sites/10^6 bp) (19). It should also be emphasised that the low level of intra-laboratory variation (corresponding to variation in procedures on different days of analysis) suggests that it may not be relevant to adhere more strictly to standard operating procedures in most of the participating laboratories, other than is already undertaken. However, the inter-laboratory variation can be reduced by standardisation of the comet assay as shown in a recent inter-laboratory validation study of the comet assay, where the inter-laboratory variation of FPG-sensitive sites was decreased when samples were analysed with a standardised comet assay protocol instead of laboratory-specific protocols (19).

The contribution of the inter-laboratory and intra-laboratory variations to the overall variation of primary comet assay endpoints was also assessed in the calibration curve data from the three different days analysed by 11 different laboratories. In this model, the overall variation was to a large extent explained by the dose of γ-radiation (60.9%), followed by the inter-laboratory variation (19.4%; $P < 0.001$ for both variables), the intra-laboratory variation (0.1%) and the unexplained variation (19.5%). We have previously made a similar assessment by comparing the calibration curves from three different ECVAG studies (16), and although the dose range of the calibration curves was narrower in this study (0–7.5 Gy compared with 0–10 Gy in the previous study), the results were similar. As in the previous ECVAG studies (16, 17), the variation of the reported primary comet assay endpoints was large (ranging from 13.9 to 78.9 %DNA in tail and a.u. for the samples that had been exposed to 7.5 Gy; Figure 1A). This variation is likely to be explained both by differences in image analysis (and possibly staining) (16) and differences in the used comet assay protocols (17). Although the analyses did not have precisely the same design, the contribution of the dose of the γ-radiation to the overall variation, 60.9%, can be compared with the contribution of the sample in the analysis of the PBMC, which accounted for <1% of the overall variation (Figure 2). The large difference in outcome of these analyses is explained by the range of damage in the sample sets, with a large range in the calibration curve samples and a narrow range in the PBMC samples, where the three donors were all young healthy women with relatively low and similar levels of DNA damage (Table IV).

Only half of the participating laboratories found a significant dose–response relationship in the coded calibration curve samples on all three analysis occasions (Figure 1). This reflects a poorer performance in this study compared with two of the previous ECVAG studies, where all but one laboratory (in which the dose–response relationship was of borderline significance) (16) and all participants (17) found a significant linear dose–response relationship in the calibration curve samples. The range of doses in the calibration curve was, however, narrower in this study (0–7.5 Gy) compared with 0–10 Gy in these two previous ECVAG studies. Six of the seven laboratories that had problems in detecting a dose–response relationship in the calibration curve samples in this study also had serious problems in another recent ECVAG study (19). The problems included negative values of FPG-sensitive sites, very low levels of FPG-sensitive sites detected in a positive control, the lack of a dose–response relationship for the calibration curve or a lack of data points due to loss of gels (19). As previously pointed out by Dusinska and Collins (26), the value of the comet assay depends on how well it is learned and performed.

In conclusion, in this study, inter-laboratory variation accounted for the largest fraction of the overall variation and the unexplained (residual) variation was much larger than the intra-laboratory variation when analysing coded PBMC samples with the comet assay using a balanced Latin square design. It is still difficult to compare values of DNA damage in PBMC from different laboratories, but the low intra-laboratory variation indicates a stability of the comet assay over time in some of the participating laboratories. This supports the notion that the comet assay is a useful tool in biomonitoring and genetic toxicology in laboratories with a high level of reproducibility.

References

1. Collins A. R. (2004) The comet assay for DNA damage and repair: principles, applications, and limitations. Mol. Biotechnol., 26, 249–261.

2. Spanswick V. J., Hartley J. M., Hartley J. A. (2010) Measurement of DNA interstrand crosslinking in individual cells using the single cell gel electrophoresis (Comet) assay. Methods Mol. Biol., 613, 267–282.
3. Shaposhnikov S., Thomsen P. D., Collins A. R. (2011) Combining fluorescent in situ hybridization with the comet assay for targeted examination of DNA damage and repair. Methods Mol. Biol., 682, 115–132.
4. Collins A. R., Dusinská M., Horváthová E., Munro E., Savio M., Stětina R. (2001) Inter-individual differences in repair of DNA base oxidation, measured in vitro with the comet assay. Mutagenesis, 16, 297–301.
5. Gaivão I., Piasek A., Brevik A., Shaposhnikov S., Collins A. R. (2009) Comet assay-based methods for measuring DNA repair in vitro; estimates of inter- and intra-individual variation. Cell Biol. Toxicol., 25, 45–52.
6. Herrera M., Dominguez G., Garcia J. M, et al. (2009) Differences in repair of DNA cross-links between lymphocytes and epithelial tumor cells from colon cancer patients measured in vitro with the comet assay. Clin. Cancer Res., 15, 5466–5472.
7. Langie S. A., Wilms L. C., Hämäläinen S., Kleinjans J. C., Godschalk R. W., van Schooten F. J. (2010) Modulation of nucleotide excision repair in human lymphocytes by genetic and dietary factors. Br. J. Nutr., 103, 490–501.
8. Tice R. (1995) The single cell gel/comet assay: a microgel electrophoretic technique for the detection of DNA damage and repair in individual cells. In Phillips D. H., Venitt S . (eds), Environmental Mutagenesis. Bios Scientific Publishers, Oxford, pp. 315–339.
9. Albertini R. J., Anderson D., Douglas G. R, et al. (2000) IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans. International Programme on Chemical Safety. Mutat. Res., 463, 111–172.
10. Collins A. R. (2011) The use of bacterial repair endonucleases in the comet assay. Methods Mol. Biol., 691, 137–147.
11. Hartmann A., Agurell E., Beevers C, et al. 4th International Comet Assay Workshop. (2003) Recommendations for conducting the in vivo alkaline Comet assay. 4th International Comet Assay Workshop. Mutagenesis, 18, 45–51.
12. Rojas E., Lopez M. C., Valverde M. (1999) Single cell gel electrophoresis assay: methodology and applications. J. Chromatogr. B Biomed. Sci. Appl., 722, 225–254.
13. Tice R. R., Agurell E., Anderson D, et al. (2000) Single cell gel/comet assay: guidelines for in vitro and in vivo genetic toxicology testing. Environ. Mol. Mutagen., 35, 206–221.
14. Azqueta A., Gutzkow K. B., Brunborg G., Collins A. R. (2011) Towards a more reliable comet assay: optimising agarose concentration, unwinding time and electrophoresis conditions. Mutat. Res., 724, 41–45.
15. Ersson C., Möller L. (2011) The effects on DNA migration of altering parameters in the comet assay protocol such as agarose density, electrophoresis conditions and durations of the enzyme or the alkaline treatments. Mutagenesis, 26, 689–695.
16. Forchhammer L., Johansson C., Loft S, et al. (2010) Variation in the measurement of DNA damage by comet assay measured by the ECVAG inter-laboratory validation trial. Mutagenesis, 25, 113–123.
17. Johansson C., Møller P., Forchhammer L, et al. (2010) An ECVAG trial on assessment of oxidative damage to DNA measured by the comet assay. Mutagenesis, 25, 125–132.
18. Møller P. (2006) The alkaline comet assay: towards validation in biomonitoring of DNA damaging exposures. Basic Clin. Pharmacol. Toxicol., 98, 336–345.

19. Forchhammer L., Ersson C., Loft S, et al. (2012) Inter-laboratory variation in DNA damage using a standard comet assay protocol. Mutagenesis, 27, 665–672.
20. Møller P., Möller L., Godschalk R. W., Jones G. D. (2010) Assessment and reduction of comet assay variation in relation to DNA damage: studies from the European Comet Assay Validation Group. Mutagenesis, 25, 109–111.
21. Zainol M., Stoute J., Almeida G. M., Rapp A., Bowman K. J., Jones G. D. ; ECVAG. (2009) Introducing a true internal standard for the Comet assay to minimize intra- and inter-experiment variability in measures of DNA damage and repair. Nucleic Acids Res., 37, e150.
22. Azqueta A., Meier S., Priestley C., Gutzkow K. B., Brunborg G., Sallette J., Soussaline F., Collins A. (2011) The influence of scoring method on variability in results obtained with the comet assay. Mutagenesis, 26, 393–399.
23. Collins A., Dusinská M., Franklin M, et al. (1997) Comet assay in human biomonitoring studies: reliability, validation, and applications. Environ. Mol. Mutagen., 30, 139–146.
24. Ahnström G., Erixon K. (1981)Measurement of strand breaks by alkaline denaturation and hydroxyapatite chromatography. In Friedberg E. C., Hanawalt P. C . (eds), DNA Repair: A Laboratory Manual of Research Procedures. Marcel Dekker, New York, pp. 403–418.
25. Kohn K. W., Erickson L. C., Ewig R. A., Friedman C. A. (1976) Fractionation of DNA from mammalian cells by alkaline elution. Biochemistry, 15, 4629–4637.
26. Dusinska M., Collins A. R . (2010) DNA oxidation, antioxidant effects, and DNA repair measured with the comet assay. In Aldini G., Yeum K. J., Niki E., Russell R. M . (eds), Biomarkers for Antioxidant Defense and Oxidative Damage: Principles and Practical Applications. Wiley-Blackwell, Oxford, UK, pp. 261–282.