Free Radical Research, Volume 38 Number 11 (November 2004), pp. 1207–1214

Taylor & Francis
healthsciences

Intra-laboratory Comet Assay Sample Scoring Exercise for Determination of Formamidopyrimidine DNA Glycosylase Sites in Human Mononuclear Blood Cell DNA

PETER MØLLER*, GITTE FRIIS, PERNILLE HØGH CHRISTENSEN, LOTTE RISOM, GALINA PLESNER, JANNE KJÆRSGAARD, PETER VINZENTS, STEFFEN LOFT, ANNIE JENSEN and MICHALA TVED

Institute of Public Health (c/o Department of Pharmacology), The Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200, Copenhagen N, Denmark

Accepted by Professor A. Collins

(Received 24 June 2004; In revised form 12 October 2004)

Oxidative DNA damage detected by the comet assay as formamidopyrimidine DNA glycosylase (FPG) senstitive sites, almost as a rule is reported as comet assay score rather than numerical sites in the genome, probably because the latter requires X-ray calibration. We compared the ability of five experienced and five inexperienced comet assay investigators to detect a dose–response relationship in irradiated A549 lung epithelial cell culture samples (0, 10 Gy and three samples of 5 Gy), based on an arbitrary five class scoring system. The samples were scored on three different occasions, thus allowing determination of the variation in sample scoring. All investigators qualitatively distinguished between samples in a dose-dependent manner, albeit with large variation in the slope and intercept of dose–response curves. There was a tendency that investigators with experience in scoring A549 cells had more consistent results than experienced investigators who had only scored lymphocytes or inexperienced investigators. The inexperienced investigators improved their scoring ability during the three sessions. Subsequently we showed that the variation in baseline level of FPG modifications in mononuclear blood cells of five healthy humans was lower when investigators used their individual X-ray calibration curve as compared to a common calibration curve. In conclusion, this study showed that comet assay investigators score differently when using a five class scoring system, which indicates that more consistent estimations of FPG sites in the genome are obtained by use of investigators' individual X-ray calibrations.

Keywords: Comet assay; FPG protein; Variation; Oxidative DNA damage

INTRODUCTION

In recent years there has been much concern over the real level of oxidative DNA damage in animal or human tissues or cell cultures. Based on inter-laboratory validation of 7-hydro-8-oxo-2′-deoxyguanosine (8-oxodG) measurements by chemical or enzymic detection, the European Standards Committee on Oxidative DNA Damage (ESCODD) has concluded that the true level of 8-oxodG in cells probably is 0.3–4.2 modifications/10^6 dG.[1] However this consensus level somewhat disguises the discrepancy between chemical end enzymic detection of 8-oxodG that is in the excess of 5-fold.[2] One reason for the lower estimation of oxidative DNA damage by the enzymic assays could be because the majority of the results was obtained by the FPG-modified version of the comet assay that have used the same X-ray calibration curve, although it justifiably should be emphasized that enzymic detection of 8-oxodG by the FPG-modified alkaline elution and alkaline unwinding assays provide similar results as the comet assay even though they are independently calibrated. The aims of this study were to investigate the differences in sample scoring by the comet assay, and to estimate the variation of FPG sensitive sites in mononuclear blood cells (MNBC) of humans when investigators use their individual X-ray calibration curve. Our laboratory experience with the arbitrary five class scoring system

*Corresponding author. Tel.: +45-35327654. Fax: +45-35327610. E-mail: p.moller@pubhealth.ku.dk

ISSN 1071-5762 print/ISSN 1029-2470 online © 2004 Taylor & Francis Ltd
DOI: 10.1080/10715760400016824

shows that inexperienced comet assay investigators can detect qualitative differences in a dose–response, dependent manner, but usually have day-to-day variation that could be due to both variation in procedure and sample scoring. Also we have noted that highly experienced comet assay investigators detect dose–response relationships of DNA strand breaking agents, although they produce quite different scores. In this investigation we first studied the ability of experienced and inexperienced comet assay investigators to detect qualitative differences between samples irradiated with X-rays and estimated the variation in sample scoring. In the second part of the study we assessed the variation in determination of FPG sensitive sites in MNBC when investigators used individual X-ray calibration curves.

METHODS

Investigators

Five investigators with experience of scoring comet assay samples participated in the study. Three of these had experience with scoring samples of cell cultures or animal tissues (investigators 2, 3 and 5), and two had experience only with scoring samples of MNBC in biomonitoring studies (investigators 1 and 4). Investigator 4 was classified as experienced although the person had not scored samples in the last two years before the experiment. From the department staff we selected five investigators (investigators 6–10) who had general knowledge of the comet assay, yet had never scored samples. The inexperienced investigators had a 15 min introduction to the microscope and sample scoring before the first occasion they scored samples. This included a description of the five class scoring system set up by Collins *et al.*[3] We used the same set of illustrations of typical comets that were distributed to ESCODD members in the inter-laboratory standardization exercises. All samples were processed through the comet assay procedure by investigator 2, and all investigators scored the samples prepared by investigator 2.

Assessment of Sample Scoring Variation

The first part of the study aimed to investigate the feasibility of a simple sample scoring exercise with all investigators scoring the same set of samples. Human A549 lung epithelial cells (from American Type Culture Collection) were embedded in 0.75% low melting point agarose on Gelbond films (BioWhittaker, Molecular Applications, Rockland, ME, USA) and irradiated with X-rays at 0, 5 and 10 Gy using a Stabilipan (Siemens) therapeutic unit (dose rate of 4.68 Gy/min at 300 kV and 12 mA). Six samples were irradiated on Gelbond films per radiation dose and all samples for the same dose should thus have received the same level of X-rays, i.e. the variation contributed by different X-ray irradiations could be ignored.

The comet assay was performed as reported previously.[4] Each investigator scored 100 images per sample (range in comet score was 0–400 arbitrary units). The investigators were instructed to score five coded samples, comprising cells exposed to 0, 5 (three samples) and 10 Gy. Preliminary experiments had indicated that samples irradiated with these doses of X-rays were sufficiently diverse to allow experienced investigators to clearly distinguish them in a dose-dependent manner. In our experience, these doses of X-rays produced samples with homogenous appearing images (the samples mainly falling into class 0, 2 or 3 and 3 or 4 for the 0, 5 and 10 Gy irradiated cells, respectively). On three different occasions the investigators scored the five samples unaware that these were identical. Eight of the investigators scored the samples on separate days, whereas investigators 5 and 8 had two sample scoring sessions on one day (separated as scoring sessions in the morning and late afternoon). The samples were re-stained three times during the scoring when the images were beginning to fade (approximately every second day, corresponding to 16 h of analysis at the microscope). The code of the samples was changed before the second and third occasion of scoring. Using this design, we are able to calculate the variance between samples and the day-to-day variance (i.e. the variance related to the different occasions of scoring are referred to as day-to-day variance). Especially the latter was of major concern because it could not be ruled out beforehand that the images would appear different because of the extensive scoring and re-staining, or because the lamp in the fluorescence microscope might have dimmed.

The performance of the investigators was evaluated on both a qualitative scale (as the variance of samples irradiated with 5 Gy) and a quantitative scale (as values of correlation coefficients of linear dose–response relationship of samples irradiated with 0, 5 and 10 Gy). For the quantitative assessment, we calculated the total variance of samples irradiated with 5 Gy (nine samples for each investigator) and determined the contribution of sample variance (SS_{sample}) and day-to-day variance (SS_{day}) by two-factor ANOVA test for single factor effects. We used three 5 Gy samples to achieve a balanced design with three occasions of scoring (3×3 design with one observation in each cell), because the heterogeneity in the measurement is distributed equally to the factors in this design (i.e. day and sample). Because of the statistical design with one observation in each group, it is not possible to test for interactions

between samples and day-to-day differences in sample scoring. The data are outlined as sum of squares (SS) below. Theoretically, the range of SS can maximally vary between zero (the score of all nine samples are the same) and 355556 (four samples with score 0 and 5 samples with score 400). The qualitative performance of the investigators was assessed by the dose–response curves with correlation coefficients (r) calculated on each sample scoring occasion (the mean of the 5 Gy irradiated samples was calculated, i.e. the datasets contained three values for 0, 5 and 10 Gy). The correlation coefficients were calculated for each scoring occasion and the three occasions combined.

Assessment of FPG Sites in MNBC

The second part of the study was initiated to determine the variation of the number of FPG sites in MNBC of healthy human subjects when investigators used individual or common X-ray calibration curves for the conversion of comet score to numerical FPG sites per diploid cell or unaltered bases. Venous MNBC were isolated from five healthy subjects by the Lymphoprep isolation procedure as described by the manufacturer (Axis-Shield PoC AS, Norway), and the cells were embedded in 0.75% agarose on Gelbond films. The level of FPG sensitive sites was obtained as the difference in score between samples incubated with buffer or FPG enzyme (1 μg/ml, kind gift of professor Andrew Collins, University of Oslo, Norway) for 45 min as described previously.[5] All investigators scored the same samples of MNBC (one sample for each donor). The samples were re-stained once after half of the investigators had scored the samples because they were beginning to fade. The calibration curves consisted of three samples of X-ray irradiated A549 cells (0, 2 and 5 Gy) that were scored on the same occasion as the investigators scored the MNBC samples. The same set of A549 cell samples and MNBC were scored by all the investigators. The 0–5 Gy dose-interval was chosen because the score of FPG sensitive sites in MNBC is considerably lower than the level of damage induced by 10 Gy X-ray. Statistically significant differences in the level of DNA damage between MNBC samples, and differences in scoring by different investigators, were analyzed by a two-factor ANOVA for single factor effects of MNBC samples and investigators.

The number of FPG sites in MNBC was calculated according to the estimations of the yield of strand breaks made by X-ray irradiation in cell culture experiments as described previously.[4] The average strand break yield per Gy has been estimated to 0.27 and 0.31 breaks/10^9 Dalton using alkaline sucrose sedimentation technique.[6,7] Using the average of the two estimations (0.29 breaks/10^9 Dalton), this corresponds to 0.19 modifications/10^6 bp per Gy (or 1160 modifications per diploid cell per Gy), assuming that mammalian cells contain 4×10^{12} Dalton DNA or 6×10^9 bp. From the comet score of MNBC, the corresponding level of Gy-equivalents was calculated using the slope of the individual X-ray calibration curve. The corresponding number of FPG sites was calculated from Gy-equivalents using the conversion factor, i.e. 0.19 modifications/10^6 bp per Gy (or 1160 modifications/diploid cell per Gy). We used the ESCODD X-ray dose–response curve with a slope of 0.0238 Gy/score as an example of a calibration curve common to all investigators.

RESULTS AND DISCUSSION

Assessment of Sample Scoring Variation

We observed differences in the comet score (in arbitrary units) between the investigators (Table I). The mean (SD) score of the samples irradiated with 5 Gy was 163 (61). In general, the coefficient of variation (CV) for each investigator was below 20%, with two exceptions (investigator 1 and 4). This indicates that the CVs generally were lower than CVs reported for the whole comet assay procedure, i.e. 18,[8] 23–38,[5] 36,[9] 42,[10] 42,[11] 46[12] and 33–95%.[13] Also the CVs are lower than those found for control groups in biomonitoring

TABLE I Statistics of data from samples irradiated with 5 Gy of X-rays[a] and coefficients of variation (in percent) for samples irradiated with 0 and 10 Gy

Investigator	Comet score[b]	SS_{total}	SS_{day}	SS_{sample}	CV (0 Gy)	CV (10 Gy)
1	32 ± 16 (50)	2071	214	1654*	100	4.1
2	120 ± 9 (8)	672	361	41	45	0.4
3	154 ± 15 (10)	1794	1568*	109	33	10.3
4	220 ± 97 (44)	75066	69624*	5299*	84	14.6
5	231 ± 32 (14)	7938	4270	2222	33	2.9
6	112 ± 20 (18)	3216	2130*	636	7	25.6
7	177 ± 20 (15)	5576	2293	950	96	4.7
8	182 ± 30 (16)	7107	2720	2456	90	3.5
9	192 ± 33 (17)	8718	5620	720	85	17.9
10	214 ± 42 (20)	14067	8623	1070	76	4.7

[a]The total variance are expressed as the sum of squares (SS) with break down into contributions of the day (SS_{day}) and sample (SS_{sample}). The residual variance can be calculated by subtraction of SS_{day} and SS_{sample} from SS_{total}. Investigators 1–5 are experienced and investigators 6–10 are inexperienced in comet assay sample scoring. [b]The data are mean ± SD (CV). *Statistically significant effect ($p < 0.05$, two-factor ANOVA).

studies, i.e. the mean CV in control groups has been reported to be 36% (95% confidence interval: 27–46%).[14] We have estimated that the intra-assay and inter-assay variation of the comet assay contribute with 35 and 65% of the total variation, respectively.[12] Accordingly, with CVs less than 50% as reported in most biomonitoring studies and with 35% of the assay variation contributed by intra-assay variation, we should expect CVs less than 20% in this experiment. This also indicates that a large proportion of the intra-assay variation is contributed by variation in sample scoring.

The results did not indicate that experienced investigators as a group scored differently compared to inexperienced. The CVs for the inexperienced investigators were remarkably similar, whereas the experienced investigators 2, 3 and 5 had low CVs and investigators 1 and 4 had high CVs. Experience with particular kinds of material might influence performance, since (a) different cell types yield comets that can differ in appearance (cultured and animal tissue cells giving 'fuzzy' comet heads compared with the round heads seen with lymphocytes, for instance); and (b) comets seen in the context of biomonitoring generally show a limited range of basal levels of damage, while more severely damaged comets are the norm in cell culture or animal experiments, where damage is induced by genotoxic agents. The CVs of the 0 Gy samples were generally higher than the CVs of the 10 Gy samples (Table I). As can also be seen from Table I, the day-to-day variance (SS_{day}) was larger than the sample variance (SS_{sample}) for 9 of the 10 investigators. This indicates a larger variation in sample scoring from one day to another as compared to samples scored on the same day. Four out of the 5 statistically significant differences were seen in the group of experienced investigators. Three investigators found a statistically significant effect of the day, whereas two investigators found statistically significant differences between the samples. However, there was not a general tendency toward systematic differences in the scores of samples or day observed by the majority of the investigators.

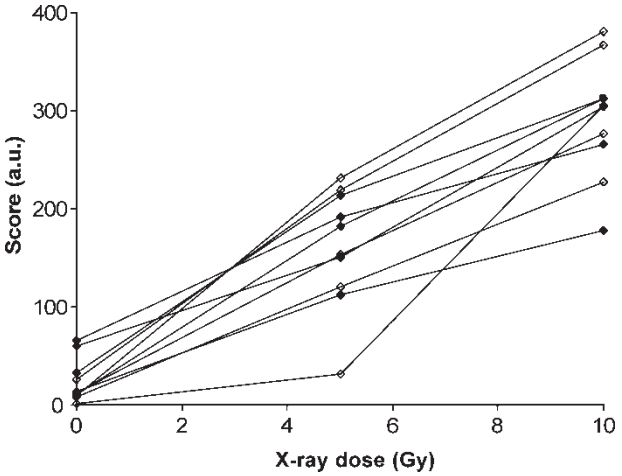

FIGURE 1 X-ray dose–response curve for experienced (open symbols) and inexperienced (solid symbols) investigators.

The dose–response relationships are presented in Fig. 1, and the slopes and intercepts are outlined in Table II. All of the investigators were able to distinguish between samples irradiated with 0, 5 and 10 Gy on every sample-scoring occasion. Although the inexperienced investigators had higher intercepts than the experienced investigators, this was only of borderline significance ($p = 0.07$, t-test). The correlation coefficients were calculated for each day of scoring and the three days combined (Table II). There was no difference in the performance between the experienced and inexperienced investigators with regard to the dose–response relationship, assessed as the magnitude of the correlation coefficients. However, a stratification of the investigators indicated that those who had experience with scoring A549 cells had the highest correlation coefficients (the mean correlation coefficients (SD) were 0.99 (0.01), 0.84 (0.02) and 0.90 (0.06) for investigators with or without A549 sample scoring experience, and inexperienced investigators, respectively). Interestingly, as can be seen from Table II there was a clear improvement in dose–response determination among the inexperienced investi-

TABLE II Slope, intercept and correlation coefficient of X-ray dose–response relationship

Investigator	Slope	Intercept	Correlation coefficient (r^2)[a]	Correlation coefficient (r^2) for each day[b]
1	6.2	1.0	0.82	0.83 (0.84;0.83;0.81)
2	22.0	9.0	1.00	1.00 (1.00;1.00;1.00)
3	26.5	15.0	0.98	0.99 (0.99;0.99;0.99)
4	34.0	34.3	0.85	0.96 (0.96;0.94;0.98)
5	37.1	21.4	0.98	0.98 (0.97;0.99;1.00)
6	16.5	19.1	0.88	0.94 (0.87;0.95;0.98)
7	24.4	49.5	0.89	0.95 (0.90;0.95;1.00)
8	30.3	17.1	0.99	0.99 (0.97;1.00;1.00)
9	20.0	74.8	0.81	0.96 (0.88;0.99;1.00)
10	27.9	46.3	0.94	0.96 (0.94;1.00;0.96)

[a]The correlation coefficient of linear regression is calculated from data of all the three sample scoring occasions (i.e. three values of 0, 5 and 10 Gy), and corresponds to the slope and intercept indicated in the table. [b]The correlation coefficient of linear regression is calculated for each sample scoring occasion, and the mean of three correlation coefficients are calculated (correlation coefficients for each day are shown in brackets).

gators, whereas this was not observed among the experienced ($p < 0.01$ for the interaction between group and repeats in repeated measurements ANOVA). This presumably reflects the learning process of scoring comets.

In order to rank the investigator performance, we made a combined ranking based on the sum of the ranking in qualitative and quantitative sample scoring exercises (low values are equal to best ranking). The qualitative and quantitative rankings were based on the values of the total variance (SS_{total}) shown in Table I, and correlation coefficient (r^2) shown in Table II, respectively. The overall ranking, based on the performance in qualitative and quantitative ranking, revealed a slightly better ranking among the experienced investigators (rank 4.7 versus 6.3 for the experienced and inexperienced, respectively). However, a stratification of the experienced investigators indicated that those with prior knowledge of scoring A549 cell samples had better ranking than investigators without prior knowledge (rank 2.3 versus 8.3 for the investigators with or without prior knowledge of scoring A549 samples, respectively). Although this suggests that experience matters when scoring comets, there is an unfortunate coincidence between the investigators with the best ranking and those who initiated and enforced the progress of the study. Considering this also was an investigation of the feasibility of carrying out multiple scorings of identical samples, several logistic considerations and issues related to the stability of the samples had to be made during the study. It would not have been possible for someone without detailed knowledge of the comet assay procedure to manage the progress of the experiments, and the statistical power of the study would have been dramatically lower if one of the experienced investigators had been excluded for the sake of managing the study.

Assessment of FPG Sites in MNBC

The second part of the study was initiated with the aim of assessing the level of FPG sites in MNBC of healthy human subjects by conversion of the comet score to numerical sites per unaltered bases. As observed in the first experiment, the slopes and intercepts of the X-ray dose–response curves differed considerably between the different investigators (Table III). Two of the investigators obtained negative results of FPG sensitive sites in MNBC and these results were considered as zero score (one and two samples scored by investigator 9 and 5, respectively). The results of comet score and FPG sites calculated using the investigator's individual and the common calibration curve are outlined in Fig. 2. Statistically significant differences of the MNBC samples and investigators were tested by two-factor ANOVA for single factor effects. This showed that the MNBC samples had different levels of damage expressed as comet score or numerical FPG sites calculated by a common calibration curve ($p < 0.01$, single factor effect of sample), whereas the effect of the investigator was of borderline significance ($p = 0.06$, single factor effect of investigator). For the numerical FPG sites calculated by the individual calibration curve, the statistical analysis revealed an effect of the sample ($p < 0.01$, single factor effect of sample), whereas the effect of the investigator was not statistically significant ($p = 0.34$, single factor effect of investigator). The majority of the investigators found a higher level of oxidative DNA damage, expressed as comet score or FPG modifications/10^6 dG, in the sample obtained from subject number 4. Although not backed up by statistical significance at 5% level, the investigators appeared to differ in their comet scores (Table IV). The mean of the five samples ranged from 11 (investigator with lowest mean score) to 54 (investigator with highest mean score). The mean (SD) score was 23 (14) which corresponds to 59% CV. By using a common X-ray calibration curve, exemplified by the ESCODD calibration curve, the mean (SD) score was 0.25 (0.15) FPG sites/10^6 dG (0.12 and 0.56 FPG sites/10^6 dG for the investigators with the lowest and highest values, respectively). When the level of FPG modifications was calculated with the investigators individual X-ray calibration curve, the mean (SD) was 0.26 (0.10) FPG sites/10^6 dG (0.09 and 0.49 FPG sites/10^6 dG for the investigators with lowest and highest values, respectively). This corresponded to a 39% CV, which is lower than that obtained by using the common X-ray calibration curve, although the variances were not statistically different ($p > 0.05$, parametric test for difference in variance between two groups). The critical CV for samples calculated by the ESCODD X-ray calibration curve is 88% for statistical significance at 5% level. It is worthwhile to point out that the group size

TABLE III Slope and intercept of X-ray dose–response curves used as individual calibration curves[a]

Investigator	Slope	Intercept	Correlation coefficients (r^2)
1	17.9	15.2	0.99
2	22.4	17.7	0.95
3	31.9	5.6	0.99
4	63.3	−0.7	0.94
5	54.6	2.3	0.99
6	27.3	14.9	0.99
7	33.9	68.8	0.96
8	51.9	35.0	1.0
9	37.0	32.0	1.0
10	48.7	18.9	0.99

[a]The dose–response curves are obtained from datasets of 0, 2 and 5 Gy. Correlation coefficients correspond to linear regression. Investigators 1–5 are experienced and investigators 6–10 are inexperienced in comet assay sample scoring.

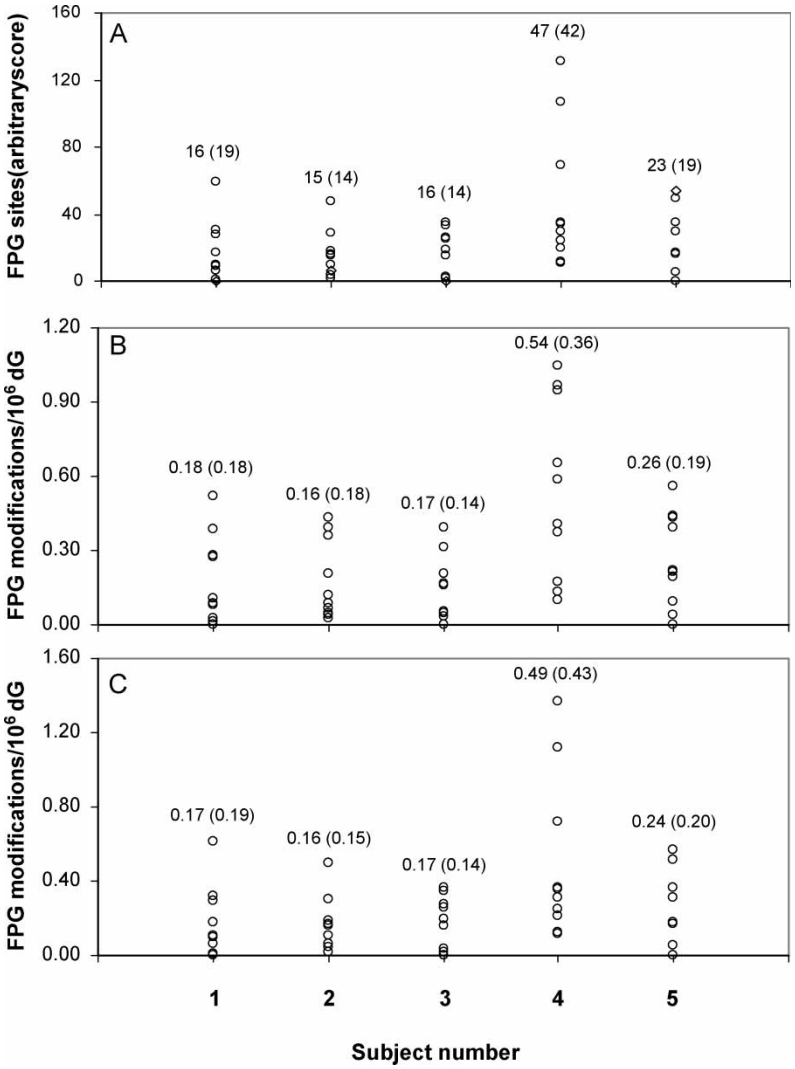

FIGURE 2 Level of FPG modifications in MNBC of healthy human subjects outlined as comet score (A) or numerical FPG values estimated by individual (B) or common (C) calibration curves. The common calibration is identical to the calibration that was used in the ESCODD project. Test for single factor effect show statistically significant effect of the subject ($p < 0.01$, two-factor ANOVA), whereas the effect of investigator is statistically insignificant. Mean (SD) values for each subject is indicated above the columns.

TABLE IV Level of FPG sites in MNBC of humans[a]

		FPG sites/10^6 dG	
Investigator	Score	Common calibration	Individual calibration
1	12 (10)	0.12 (0.10)	0.29 (0.24)
2	15 (16)	0.15 (0.16)	0.27 (0.29)
3	24 (25)	0.25 (0.27)	0.33 (0.35)
4	31 (17)	0.32 (0.17)	0.20 (0.11)
5	38 (57)	0.40 (0.59)	0.31 (0.45)
6	14 (13)	0.14 (0.14)	0.22 (0.21)
7	20 (12)	0.21 (0.13)	0.25 (0.15)
8	11 (5)	0.12 (0.05)	0.09 (0.04)
9	16 (17)	0.16 (0.17)	0.19 (0.20)
10	54 (31)	0.56 (0.32)	0.49 (0.28)
Mean ± SD (CV)	23 ± 14 (59%)	0.25 ± 0.15 (59%)	0.26 ± 0.10 (39%)

[a]The data are expressed as the mean (SD) of five MNBC samples for each investigator. Score represents the comet assay score in arbitrary units, whereas the number of FPG sites/10^6 dG is calculated by use of the common and individual calibration curve.

TABLE V Assessment of FPG lesions in MNBC[a]

Study	Original reported	FPG sites/10^6 dG	FPG sites/cell
Møller *et al.*, 2004[15]	12.5 arbitrary units	0.17[b]	452[b]
This study		0.26	697
Collins *et al.*, 1997[16]	870 sites/cell	0.33	870
ESCODD, 2004[1]	0.34 sites/10^6 dG	0.34 (0.15)[c]	898 (396)[c]
Møller *et al.*, 2003[5]	37.1 arbitrary units	0.39[d]	1030[d]
Collins *et al.*, 1996[17]	0.28 sites/10^9 Dalton	0.42	1100
Merzenich *et al.*, 2001[18]	0.23 sites/10^6 bp	0.52	1375
Pfaum *et al.*, 1997[19]	0.24 sites/10^6 bp	0.55	1440
Møller *et al.*, 2004[12]	49.8 arbitrary units	0.68[b]	1802[b]
Gedik *et al.*, 2002[20]	1.33 sites/10^6 dG	1.33	3511

[a]The calculation of the number of FPG modifications is based on assumptions that human diploid cells contain 4×10^{12} Dalton DNA (corresponding to 6×10^9 bp), and that 22% of the bases are guanines. [b]The calculation is based on the individual X-ray calibration curve used in this study. [c]The median value of FPG sites reported by different laboratories in the ESCODD study is shown, with in brackets the number of FPG sites in genomic DNA reported from this laboratory. [d]The calculation is based on the ESCODD calibration curve.

(5 MNBC samples) in this experiment is considerably lower than the number of subjects normally recruited in biomonitoring studies, and a group size comparable to that used in biomonitoring studies probably would have yielded statistically significant effects of the investigator. An overall interpretation of these results suggests that the advantage of using individual calibration curves is related to the consistency of the results (i.e. less variance between investigators), whereas the common and individual calibration curves produce remarkably similar estimates of the number of FPG sites.

Our estimate of the level of FPG modifications is in the lower end of the reported estimations (Table V). The ESCODD partners have concluded that the background level of 8-oxodG in genomic DNA is likely to be in the range of 0.3–4.2 lesions/10^6 dG (790–11100 lesions/diploid cell), based on both evaluation of enzymic and chemical detection of oxidative DNA damage in MNBC samples.[1] The lower level of oxidative guanine modifications (790 FPG modifications/diploid cell) is close to our estimation of the background level of oxidative DNA damage in MNBC, i.e. 251–1284 FPG sites/diploid cell for investigators producing the lowest and highest mean of the five MNBC samples.

Curiously, all the investigators produced X-ray calibration curves with larger slopes in the second part of the study compared to the first part (mean (SD) was 38.9 (15) score/Gy and 24.5 (9.0) score/Gy for the second and first part of the study, respectively ($p < 0.001$, paired t-test)). This is striking because the 0 and 5 Gy samples were identical in the first and second X-ray dose–response curves. It is possible that the samples irradiated with 5 Gy in the second part of the study had slightly higher damage level despite that they were irradiated at the same time. In this case, the sample variation would be much larger than estimated in the first part of the study, since the mean comet score of all investigators was 218 in the last experiment and 160, 164, 167 in the first experiment. It is possible that the visual interpretation of comet images depend on the context in which they are scored by the investigator, i.e. the 5 Gy samples were in the middle of the dose–response curve in the first part of the study, whereas it was the most damaged in the second part of the study. As the most important outcome of this discrepancy, the estimation of FPG modifications in

TABLE VI Estimation of background FPG sites in A549 lung epithelial cells[a]

Investigator	0–10 Gy dose–response curve		0–5 Gy dose–response curve	
	FPG/10^6 dG	FPG/diploid cell	FPG/10^6 dG	FPG/diploid cell
1	0.07	184	0.41	1081
2	0.16	431	0.17	458
3	0.20	516	0.14	357
4	0.34	894	0.19	504
5	0.11	286	0.09	230
6	0.37	969	0.17	458
7	1.06	2807	1.02	2687
8	0.15	388	0.27	725
9	1.43	3785	0.43	894
10	0.50	1318	0.16	421
Mean (SD)	0.44 (0.46)	1158 (1202)	0.30 (0.27)	785 (718)

[a]The calculations are based on the scores obtained in cell samples that were not irradiated with X-rays (0 Gy). The first investigation corresponds to the 0–10 Gy X-ray dose–response curves, and the second investigation corresponds to the 0–5 Gy X-ray dose–response curve. Investigators 1–5 are experienced and investigators 6–10 are inexperienced in comet assay sample scoring.

MNBC is higher when using the individual X-ray calibration of the first part of the study with 0, 5 and 10 Gy irradiated cells (0.34 FPG sites/10^6 dG or 909 sites/diploid cell) compared to calculations made in the second part of the study with 0, 2 and 5 Gy irradiated samples for the X-ray calibration curve (0.26 FPG sites/10^6 dG or 697 sites/diploid cell). A similar tendency was seen for the background level of FPG sites in A549 cells (Table VI). This may indicate that use of X-ray calibration curves with high doses will slightly overestimate the level of oxidative modifications.

The subjectivity in the scoring of nuclei may be interpreted as a problem of the arbitrary comet score analysis. In fact, it is possible that a large proportion of the intra-assay variation may be due to the scoring of nuclei. It is unresolved, whether or not subjectivity in the scoring of nuclei is overcome by use of image analysis systems. In our experience (with the Kinetic Imaging comet analysis system) subjectivity also is required in image analysis systems. The finding highlights the need for randomisation of samples and blind scoring. It is possible that the variation due to subjectivity is reduced if the number of nuclei scored per sample is increased. Also employing fully automated comet assay image analysis systems, without the possibility of human intervention, would eliminate the subjectivity of the scoring. It should also be put into perspective that the best-trained investigators in the study had the least variation in the scoring, and it is trained personnel who routinely run comet analysis. To fully outline the role of subjectivity in sample scoring, future investigations should include only trained personnel and the design should be optimized for the purpose of testing this effect.

In conclusion, this study has shown that comet assay investigators differ in their analysis of comet samples using visual scoring. Use of individual or common (ESCODD) X-ray calibration curves produced remarkably similar estimations of the number of FPG lesions in MNBC, whereas the least variation in the number of FPG sites was achieved by using the investigators' individual X-ray calibration curves. This supports the use of individual X-ray calibration curves if calculations of numerical FPG sites in the genome are attempted.

References

[1] Collins, A.R., Cadet, J., Moller, L., Poulsen, H.E. and Vina, J. (2004) "Are we sure we know how to measure 8-oxo-7,8-dihydroguanine in DNA from human cells?", *Arch. Biochem. Biophys.* **423**, 57–65.

[2] ESCODD (European Standards Committee on Oxidative DNA Damage)(2003) "Comparative analysis of baseline 8-oxo-7,8-dihydroguanine in mammalian cell DNA, by different methods in different laboratories: an approach to consensus", *Carcinogenesis* **23**, 2129–2133.

[3] Collins, A.R., Duthie, S.J. and Dobson, V.L. (1993) "Direct enzymic detection of endogenous oxidative base damage in human lymphocyte DNA", *Carcinogenesis* **14**, 1733–1735.

[4] Risom, L., Møller, P., Vogel, U., Kristjansen, P.E.G. and Loft, S. (2003) "X-ray-induced oxidative stress: DNA damage and gene expression of HO-1, ERCC1 and OGG1 in mouse lung", *Free Radic. Res.* **37**, 957–966.

[5] Møller, P., Vogel, U., Pedersen, A., Dragsted, L.O., Sandström, B. and Loft, S. (2003) "No effect of 600 g fruit and vegetables per day on oxidative DNA damage and repair in healthy human non-smokers", *Cancer Epidemiol. Biomarkers Prev.* **12**, 1016–1022.

[6] Kohn, K.W., Erickson, L.C., Ewig, R.A.G. and Friedman, C.A. (1976) "Fractionation of DNA from mammalian cells by alkaline elution", *Biochemistry* **15**, 4629–4637.

[7] Ahnström, G. and Erixon, K. (1981) "Measurement of strand breaks by alkaline denaturation and hydroxyapatite chromatography", In: Friedberg, E.C. and Hanawalt, P.C., eds, *DNA repair: A Laboratory Manual Research Procedures* (Marcel Dekker, New York), pp 403–418.

[8] Holz, O., Jörres, R., Kästner, A. and Magnussen, H. (1995) "Differences in basal and induced DNA single-strand breaks between human peripheral monocytes and lymphocytes", *Mutat. Res.* **332**, 55–62.

[9] Speit, G., Witton-Davies, T., Heepchantree, W., Trenz, K. and Hoffmann, H. (2003) "Investigations on the effect of cigarette smoking in the comet assay", *Mutat. Res.* **542**, 33–42.

[10] Hellman, B., Friis, L., Vaghef, H. and Edling, C. (1999) "Alkaline single cell gel electrophoresis and human biomonitoring for genotoxicity: a study on subjects with residential exposure to radon", *Mutat. Res.* **442**, 121–132.

[11] Møller, P., Wallin, H., Holst, E. and Knudsen, L.E. (2002) "Sunlight induced DNA damage in human mononuclear cells", *FASEB J.* **16**, 45–53.

[12] Møller, P., Viscovich, M., Lykkefeldt, J., Loft, S., Jensen, A. and Poulsen, H.E. (2004) "Vitamin C supplementation decreases oxidative DNA damage in mononuclear blood cells of smokers", *Eur. J. Nutr.* **43**, 267–274.

[13] de Boeck, M., Hoet, P., Lombaert, N., Nemery, B., Kirsch-Volders, M. and Lison, D. (2003) "*In vivo* genotoxicity of hard metal dust: induction of micronuclei in rat type II epithelial lung cells", *Carcinogenesis* **24**, 1793–1800.

[14] Møller, P., Knudsen, L.E., Loft, S. and Wallin, H. (2000) "The comet assay as a rapid screen test in biomonitoring occupational exposure to DNA damaging agents and effect of confounding factors", *Cancer Epidemiol. Biomarkers Prev.* **9**, 1005–1015.

[15] Møller, P., Loft, S., Alfthan, G. and Freese, R. (2004) "Oxidative DNA damage in circulating mononuclear blood cells after ingestion of blackcurrant juice or anthocyanin-rich drink", *Mutat. Res.* **551**, 119–126.

[16] Collins, A.R., Duthie, S.J., Fillon, L., Gedik, C.M., Vaughan, N. and Wood, S.G. (1997) "Oxidative DNA damage in human cells: the influence of antioxidants and DNA repair", *Biochem. Soc. Trans.* **25**, 326–331.

[17] Collins, A.R., Dusinská, M., Gedik, C.M. and Stetina, R. (1996) "Oxidative damage to DNA: do we have a reliable biomarker?", *Environ. Health Perspect.* 104(Supplement 3), 465–469.

[18] Merzenich, H., Hartwig, A., Ahrens, W., Beyersmann, D., Schlepegrell, R., Scholze, M., Timm, J. and Jöckel, K.-H. (2001) "Biomonitoring on carcinogenic metals and oxidative DNA damage in a cross-sectional study", *Cancer Epidemiol. Biomarkers Prev.* **10**, 515–522.

[19] Pflaum, M., Will, O. and Epe, B. (1997) "Determination of steady-state levels of oxidative DNA base modifications in mammalian cells by means of repair endonucleases", *Carcinogenesis* **18**, 2225–2231.

[20] Gedik, C.M., Boyle, S.P., Wood, S.G., Vaughan, N.J. and Collins, A.R. (2002) "Oxidative stress in humans: validation of biomarkers of DNA damage", *Carcinogenesis* **23**, 1441–1446.