The effect of smoking on DNA effects in the comet assay: a meta-analysis

Heike Hoffmann, Josef Högel and Günter Speit*
Universität Ulm, Abteilung Humangenetik, Ulm, Germany

The comet assay (alkaline single-cell gel electrophoresis, SCG or SCGE) is frequently used in biomonitoring to detect genotoxic effects in humans exposed at the workplace or in their environment. Because of its ready accessibility, blood is most frequently used in such studies. Many studies investigated cigarette smoking either as a genotoxic exposure itself or as a potential confounding factor in occupational studies. However, although smoking is considered to be a relevant exposure towards various genotoxins, conflicting results have been reported in the comet assay studies. The actual reasons for this discrepancy are not known.

To further evaluate evidence for smoking-related DNA effects in the comet assay, we now used a meta-analysis approach based on a literature search. We identified 38 studies from 37 publications which were suited for a formal meta-analysis based on the standardized mean difference (SMD) between the study groups. The evaluation of these 38 studies indicated higher levels of DNA damage in smokers than in non-smokers [under a random effects model, SMD = 0.55, 95% confidence interval = (0.16–0.93)]. Subdividing these studies into studies investigating the effect of smoking as a genotoxic exposure (Type A studies, n = 12) and studies investigating smoking as a potential confounder in occupational studies (Type B, n = 26) indicated a significant difference only in Type A studies but not in Type B studies. Furthermore, studies using image analysis or image length measurements (n = 23) only indicated a tendency for a genotoxic effect of smoking, whereas studies using an arbitrary score (n = 15) found a significantly higher level of DNA damage in smokers.

Introduction

The comet assay (single-cell gel electrophoresis, SCG or SCGE) is a well-established genotoxicity test for in vitro and in vivo testing of chemicals (1–3). In its alkaline version, DNA strand breaks, alkali-labile sites (ALS) and incisions in the course of excision repair lead to increased DNA migration, whereas crosslinks (DNA–DNA and DNA–protein) inhibit DNA migration (4,5). Because of its ease of application and its ability to detect various kinds of DNA lesions in eukaryotic cells exposed to genotoxic agents with high sensitivity, the comet assay is also being increasingly used in human biomonitoring (for a review, see Refs 6 and 7). Within the framework of the ‘International Programme on Chemical Safety (IPCS)’ a guideline for the performance of the comet assay in human biomonitoring has been proposed (8). A specific problem of the comet assay in comparison with other standard genotoxicity tests is the fact that various protocol modifications are in use and many test variables have been identified that can influence comet assay effects (5,9). Therefore, a direct quantitative comparison of comet assay effects from various studies is generally impossible. Furthermore, evaluation of DNA effects in the comet assay can be based on visual scoring or image analysis (5). Although a visual score based on the measurement of five classes of cells with increasing damage has originally been proposed (10), other scores (e.g. based on three or four classes) are also in use. With respect to image analysis, various parameters such as ‘tail length’, ‘tail intensity’ and ‘tail moment’ can be measured and published. However, owing to the different equipment (e.g. microscope, camera) and differences in the software used (e.g. different formulas for the calculation of the tail moment) the values obtained and published cannot be directly compared. On one hand, these factors can introduce considerable inter-laboratory variability and on the other hand make comet assay effects hard to interpret. The protocol used for the comet assay in different laboratories significantly influences the sensitivity of the test. Nevertheless, many laboratories use a published standard protocol without determining and demonstrating the sensitivity of their particular protocol. These specific features of the comet assay do not cause a serious problem for appropriate in vitro and in vivo genotoxicity testing because the concentration/dose of the test compound can be increased until toxicity is reached, and positive controls can be included. However, in human biomonitoring there are characteristic limitations owing to the fact that in most cases exposure to genotoxins is only small, that there is no possibility to determine dose–response relationships, to include appropriate positive controls, or to independently reproduce the test (11). Thus the quality of comet assay results in biomonitoring particularly depends on standardized and controlled test conditions. Measures of quality assurance such as the concurrent processing and evaluation of coded and mixed slides from ‘exposed’ and ‘control’ subjects, repeated evaluation by different investigators and the inclusion of a reference standard have been proposed (8,12) but are hardly ever used.

In the course of the last years we have been critically evaluating some aspects of the comet assay with respect to its reliable use in human biomonitoring. We comparatively investigated comet assay effects in smokers and non-smokers because smoking is supposed to be a relevant exposure to genotoxins. Cigarette smoke is known to contain many mutagens/carcinogens, with polycyclic aromatic hydrocarbons (PAHs), aromatic amines, N-nitrosamines and aldehydes representing the major classes of harmful substances (13,14). Genotoxic effects of tobacco smoke have recently been reviewed (15,16) and smoking is a well-documented cause of...
cancer (17,18). The comet assay has repeatedly been used to measure DNA damage related to tobacco smoking. Some studies were performed specifically to investigate a potential effect of smoking in the comet assay, whereas many other studies determined the effect of smoking as a potential confounding factor in the course of occupational studies. Conflicting results have been reported and reviewed (6,7). Various explanations for the reported discrepancies have been proposed, including the power of the statistical analysis and also seasonal and regional differences. However, these explanations are just speculations and can hardly be verified. Therefore, to further evaluate evidence for smoking-related DNA effects in the comet assay, we now used a meta-analysis approach based on a literature search.

**Studies included in the meta-analysis**

The focus of this meta-analysis is on studies that investigate potential effects of cigarette smoking on DNA using the comet assay. Two types of investigations were eligible for this analysis which will be referred to as Type A and Type B studies. Type A, the assessment of smoking effects was the primary goal of the study that, at least, comprised a group of actual smokers and an independent control group of non-smokers. Type B, the assessment of smoking effects was a secondary goal of the study (usually an occupational study). In each study of this type, to avoid potential confounding, a group of individuals without the primary exposure had to be available that included well-defined subgroups of actual smokers and non-smokers. Candidate studies were searched for in PubMed (finalized June 30, 2005) using the terms ‘comet smok*’, ‘scg smok*’ and ‘scge smok*’. This strategy yielded a total of 167 studies including 6 studies with no connection to smoking effects in the comet assay (in four of them ‘SCG’ was used as an abbreviation for something else, and in two it was unclear why these studies were identified in PubMed). Of the remaining 161 studies (6,7,19 – 177) a number of studies had to be excluded from analysis for various principal reasons:

- 11 studies: language not English (19–29);
- 5 studies: no cigarette smoking but other sources of ‘smoke’ (30–34);
- 1 study: use of the neutral comet assay (35);
- 28 studies: no human peripheral blood cells but other target cells or in vitro studies (36–63);
- 1 study: investigation of only one subject (64);
- 18 studies: no concurrent consideration of adequate smoker and non-smoker groups [only non-smokers (8), only smokers (1), only ex-smokers (2), only passive smokers (1), smoking only in combination with drinking (1), smoking only in combination with the induction of DNA damage (5)] (65–82);
- 17 studies: no differentiation between smokers and non-smokers in the control group of Type B studies (83–99);
- 16 studies: complete absence of data or of descriptive statistical information in the control group of these Type B studies (100–115);
- 8 studies: review without new data (6,7,116–121).

Furthermore, only studies that investigated at least one of the following quantitative outcome variables were taken into consideration: tail moment (including ‘Olive tail moment’), tail intensity (including ‘percent DNA in tail’), tail length, tail migration, comet length or an arbitrary score calculated from cells in different damage classes [‘visual score’, ‘genetic damage index’ (GDI) or ‘tail factor’]. For these variables, when raw data were not given in the publications, the number of subjects, means and standard deviations in the smoker and non-smoker group were required. From the 56 studies remaining, 19 further studies failed these conditions. Fourteen of the latter studies only reported the result of a statistical test (P-value) for appropriate group comparison (122–135). These studies are summarized in a separate table (Table I) and will be considered in the discussion. Five studies stated in the text that they found a difference between smokers and non-smokers or not, but did neither show the data nor indicate a P-value (136–140). These studies could not be included in our meta-analysis but will also be considered in the discussion. Thus in total, from 161 initial candidate studies, 37 were available for a formal meta-analysis and are summarized in Table II (141–177). One publication appears twice (164) because within this paper two populations (foundry and pottery workers) were tested with their respective different controls (smokers and non-smokers). Thus our meta-analysis is based on 38 ‘studies’ from 37 publications.

**Statistical analysis**

Relevant data of the 38 studies (from 37 publications) available for formal meta-analytic evaluation was entered into the review manager ‘RevMan 4.2’, a software for documentation and evaluation of meta-analyses. It is the Cochrane Collaboration’s program for preparing and maintaining Cochrane reviews (178). For each study, the core information comprised sample sizes, mean values and standard deviations of the respective outcome variables, both for the smoker and non-smoker groups. In this context, necessary supplementations were carried out for a few studies: (i) Mean values and standard deviations were calculated from the published raw data. (ii) If only the number of cells in different damage classes was presented, we calculated the score on the basis of the available information, i.e. we gave a value from 0 to 2; from 0 to 3 or from 0 to 4 according to the given Classes 3, 4 or 5 of damage (amount of DNA migration) (10). (iii) If for a variable only the standard error of the mean (SEM) was given, we computed the standard deviation of the measurements by multiplying the SEM with the square root of the number of observations the mean value was based on.

The procedures used for statistical analysis are described in detail in the user’s guide for RevMan 4.2 (179). All outcome variables considered here were continuous. Group comparisons were performed between smokers and non-smokers both based on all 38 eligible studies as well as in the subgroups of Type A and Type B studies, as described above.

Each study was represented in the meta-analysis by one outcome measure. In eight studies more than one image analysis parameter was determined. In these cases we selected the tail moment when available (7 studies) or the tail intensity when the tail moment was not presented (1 study). However, in all of these studies, the outcome (positive or negative) was the same with all image analysis parameters shown. The studies were further divided into a group where image analysis or image length measurement were used and a group where visual score determinations (including GDI and tail factor) were available.

The various outcomes for the comet assay measure the effects of smoking on different scales. To enable a joint comparison, in each individual study the standardized mean difference (SMD) between the study groups was calculated for each outcome. The SMD is the difference between the respective mean values divided by a common measure of dispersion.
which was computed from the standard deviations in the two groups (see user’s guide: ‘Hedges adjusted g’). In our analysis, positive SMD indicates higher DNA damage in smokers. Thus, each study yields dimensionless measures of difference that constitute the ‘effect sizes’ to be combined in the meta-analysis. The transformation of score values assigned to single cells resulted in a quantitative outcome measure for each proband. We checked compatibility with a normal distribution of scores among probands in several groups and did not detect major deviations. Thus, the mean values represent the distributions quite well.

The effect sizes of the individual studies were pooled both under a ‘fixed effects model’ and a ‘random effects model’, and the results of both analyses are presented. The fixed effects model assumes a common exposure effect in the studies and the results of both analyses are presented. The fixed effects model, effect sizes are not assumed to be common to all studies but to vary themselves under a normal distribution model. Here, study weights include this additional variance (DerSimonian and Laird Model) and can be quite different from those under the fixed effects model (with a tendency to ‘equal weighting’). To decide which model is more appropriate, a heterogeneity statistic is calculated where large values correspond to ‘small’ studies. Ideally, when no bias is present, the observed deviations among the studies to be caused by sampling only. In this case, studies are weighted according to the inverse variance of the effect size. Under the random effects model, effect sizes are not assumed to be common to all studies but to vary themselves under a normal distribution model. Here, study weights include this additional variance (DerSimonian and Laird Model) and can be quite different from those under the fixed effects model (with a tendency to ‘equal weighting’). To decide which model is more appropriate, a heterogeneity statistic is calculated where large values support the random effects model; a formal statistical test is also given using this statistic (small P-values are in favor of random effects).

Under both models, the weighted mean of the effect sizes (i.e. SMD) is computed, and a 95% confidence interval (CI) is determined. If such an interval does not contain the value zero, smoking has a statistically significant effect (level of significance 0.05) in the respective analysis. An adjustment for multiple testing due to various subgroup analyses was not performed.

In order to assess the impact of a potential publication bias, a funnel plot is presented including all 38 studies. This is a plot that contrasts the estimated SMD in each study to its standard error (SE), a measure of its imprecision. Large SEs, in general, correspond to ‘small’ studies. Ideally, when no bias is present, particularly small studies with effect sizes under average should be equally represented in the meta-analysis as small studies showing effects above average.

**Results**

The characteristics of the 38 studies included in a formal meta-analysis are shown in Table II; the publication of Basaran et al. (164) appears twice, because this paper comprises two Type B studies with two independent control groups, each with smokers and non-smokers. The numbers of subjects included vary substantially; the minimal group size is 3 (both smokers and non-smokers), the median number is 17 (both smokers and non-smokers) and the maximal sizes are 134 (smokers) and 213 (non-smokers). Totally, data from 803 smokers and 959 non-smokers are summarized in this analysis. Table II gives means and standard deviations for the outcome of each study, i.e. the parameter given or selected (in case of different parameters shown) which was used in the meta-analysis. The corresponding SMDs and their 95% CIs are also graphically displayed in Figure 1. The intervals differ remarkably concerning their location and size, meaning great heterogeneity among

---

**Table I. Studies which reported the result of a statistical test but could not be considered for the meta-analysis**

<table>
<thead>
<tr>
<th>Study</th>
<th>Name</th>
<th>Smokers (N)</th>
<th>Non-smokers (N)</th>
<th>Outcome</th>
<th>Statistical test</th>
<th>P-Value</th>
<th>Result in favour of</th>
</tr>
</thead>
<tbody>
<tr>
<td>122</td>
<td>Knudsen et al., 2005</td>
<td>29</td>
<td>17</td>
<td>Score</td>
<td>Mann–Whitney U-test</td>
<td>0.003</td>
<td>Smokers</td>
</tr>
<tr>
<td>123</td>
<td>Leng et al., 2004</td>
<td>19</td>
<td>31</td>
<td>Tail moment</td>
<td>Student’s t-test</td>
<td>0.19</td>
<td>—</td>
</tr>
<tr>
<td>124</td>
<td>Demirbag et al., 2005</td>
<td>11</td>
<td>31</td>
<td>Score</td>
<td>‘Bivariate correlation’</td>
<td>0.006</td>
<td>Smokers</td>
</tr>
<tr>
<td>125</td>
<td>Boyaci et al., 2004</td>
<td>11</td>
<td>19</td>
<td>Undamaged Intermediate Tailed</td>
<td>Mann–Whitney U-test</td>
<td>&lt;0.05</td>
<td>Smokers</td>
</tr>
<tr>
<td>126</td>
<td>Polasa et al., 2004</td>
<td>10</td>
<td>10</td>
<td>Comet ratio D/L</td>
<td>Duncan’s multiple range test</td>
<td>&gt;0.05</td>
<td>—</td>
</tr>
<tr>
<td>127</td>
<td>Park and Kang, 2004</td>
<td>37</td>
<td>72</td>
<td>Tail moment</td>
<td>Student’s t-test</td>
<td>&lt;0.01</td>
<td>Smokers</td>
</tr>
<tr>
<td>128</td>
<td>Schabath et al., 2003</td>
<td>5</td>
<td>72</td>
<td>Tail moment</td>
<td>Student’s t-test</td>
<td>&lt;0.001</td>
<td>Smokers</td>
</tr>
<tr>
<td>129</td>
<td>Scheepers et al., 2002</td>
<td>31</td>
<td>19</td>
<td>Score</td>
<td>Parametric or Non-parametric test of location</td>
<td>&gt;0.05</td>
<td>—</td>
</tr>
<tr>
<td>130</td>
<td>Giovannelli et al., 2002</td>
<td>12</td>
<td>24</td>
<td>Tail intensity</td>
<td>Kruskal–Wallis ²</td>
<td>0.89</td>
<td>—</td>
</tr>
<tr>
<td>131</td>
<td>Undege and Basaran, 2002</td>
<td>23</td>
<td>10</td>
<td>Score</td>
<td>χ²-test</td>
<td>&lt;0.001</td>
<td>Smokers ²</td>
</tr>
<tr>
<td>132</td>
<td>Šrám et al., 1998</td>
<td>9</td>
<td>10</td>
<td>Tail intensity</td>
<td>Mann–Whitney U-test</td>
<td>&lt;0.05</td>
<td>Non-smokers</td>
</tr>
<tr>
<td>133</td>
<td>Šrám et al., 1998</td>
<td>11</td>
<td>30</td>
<td>Tail moment</td>
<td>Mann–Whitney U-test</td>
<td>&gt;0.05</td>
<td>Smokers</td>
</tr>
<tr>
<td>134</td>
<td>Tates et al., 1996</td>
<td>?</td>
<td>?</td>
<td>Tail length</td>
<td>Student’s t-test or Mann–Whitney U-test</td>
<td>0.045</td>
<td>Non-smokers</td>
</tr>
<tr>
<td>135</td>
<td>Betti et al., 1995</td>
<td>60</td>
<td>140</td>
<td>Comet length</td>
<td>Multiple regression ²</td>
<td>&lt;0.0001</td>
<td>Smokers</td>
</tr>
</tbody>
</table>

---

- ¹Based on in-transformed data.
- ²Typographical error suspected.
- ³Comparison of current smoker, ex-smoker and never smoker.
- ⁴Due to a higher proportion of undamaged cells in non-smoker.
- ⁵Based on log-transformed data.
the studies both in the effects detected and the precision attained. Table II gives the weights attributed to the studies in a fixed effects analysis (also shown in Figure 1 as circles), assuming that the obvious non-homogeneity of effects is due to sampling only. Under this model, the resulting SMD of all studies is 0.18 (95% CI: 0.07–0.28) in units of standard deviations, the P-value from the corresponding test against an overall null effect of smoking is 0.0009. However, the heterogeneity test yields strong evidence in favor of a random model (P < 0.0001). When the latter is assumed, the estimated total SMD is 0.55 (95% CI: 0.16–0.93), the corresponding P = 0.006. The weights of the studies are also presented in Table II; they are by far more independent of the characteristics of the individual studies than under the fixed effects model. As a consequence, the largest study (156) which strongly contributes evidence in favor of a null effect, is down-weighted concerning its influence on the overall result. To summarize, although the choice between the two models remains problematic and there is an ongoing controversy among experts, both analyses lead to the same conclusion that there is probably no large effect of smoking in the comet assay with peripheral blood cells. Furthermore, the funnel plot (Figure 2) indicates that the ‘smallest’ studies tend to show effects above average and there are no studies of similar precision that show effects under average. The funnel plot presents the relationship between the SMD and its SE for all 38 studies included in the meta-analysis (dots). The dashed vertical line represents the mean SMD observed in the studies in a fixed effects model (0.18). For example, studies with low precision (i.e. high SE) that show a large smoking effect on DNA (large SMD) show up in the right lower corner of the plot. Thus the tendency towards higher levels of DNA damage in peripheral blood cells of smokers could also be partly attributed to a publication bias or to other sources of bias present in small studies.

The 38 studies considered consist of 12 Type A studies (Figure 3) and 26 Type B studies (Figure 4). Type A studies (258 smokers and 303 non-smokers; median number of smokers and non-smokers: 20/study), where the comparison of smokers and non-smokers is the main goal, mostly support the hypothesis of higher DNA damage in smokers, both under the fixed effects and random effects model (fixed: total SMD = 0.54, 95% CI from 0.36 to 0.72, P < 0.0001; random: total

---

Table II. Studies included in the meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Name</th>
<th>Outcome</th>
<th>Smoker</th>
<th>Non-smoker</th>
<th>SMD* (95% CI)</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>141</td>
<td>Betti et al., 1994</td>
<td>Comet length</td>
<td>38</td>
<td>45.82</td>
<td>5.55</td>
<td>6.22</td>
</tr>
<tr>
<td>142</td>
<td>Dhawan et al., 2001</td>
<td>Tail moment</td>
<td>31</td>
<td>20.6</td>
<td>0.7</td>
<td>1.14</td>
</tr>
<tr>
<td>143</td>
<td>Besarati-Nir et al., 2001</td>
<td>Tail moment</td>
<td>21</td>
<td>27.10</td>
<td>1347.28</td>
<td>29.25</td>
</tr>
<tr>
<td>144</td>
<td>Anderson et al., 1994</td>
<td>Score</td>
<td>6</td>
<td>8.83</td>
<td>5.34</td>
<td>6.33</td>
</tr>
<tr>
<td>145</td>
<td>Piperakis et al., 2000</td>
<td>Score</td>
<td>20</td>
<td>99</td>
<td>10.9</td>
<td>20.78</td>
</tr>
<tr>
<td>146</td>
<td>Mohankumar et al., 2002</td>
<td>Comet length</td>
<td>16</td>
<td>36.25</td>
<td>8.45</td>
<td>10.21</td>
</tr>
<tr>
<td>147</td>
<td>Speit et al., 2003</td>
<td>Tail moment</td>
<td>20</td>
<td>0.45</td>
<td>0.07</td>
<td>0.77</td>
</tr>
<tr>
<td>148</td>
<td>Diem et al., 2002</td>
<td>Tail factor</td>
<td>28</td>
<td>4.61</td>
<td>1.47</td>
<td>2.53</td>
</tr>
<tr>
<td>149</td>
<td>Engel et al., 2004</td>
<td>Tail moment</td>
<td>10</td>
<td>2.34</td>
<td>0.7</td>
<td>10.23</td>
</tr>
<tr>
<td>150</td>
<td>Ellahene et al., 2004</td>
<td>Tail moment</td>
<td>6</td>
<td>0.75</td>
<td>0.29</td>
<td>6.06</td>
</tr>
<tr>
<td>151</td>
<td>Pitarka et al., 1999</td>
<td>Comet length</td>
<td>3</td>
<td>40.31</td>
<td>10.44</td>
<td>8.38</td>
</tr>
<tr>
<td>152</td>
<td>Zhu et al., 1999</td>
<td>Tail moment</td>
<td>15</td>
<td>39.32</td>
<td>7.12</td>
<td>26.34</td>
</tr>
<tr>
<td>153</td>
<td>Palus et al., 1999</td>
<td>Score</td>
<td>18</td>
<td>13.2</td>
<td>6</td>
<td>23.7</td>
</tr>
<tr>
<td>154</td>
<td>Šárdaš et al., 1997</td>
<td>Tail migration</td>
<td>5</td>
<td>2.52</td>
<td>0.46</td>
<td>3.04</td>
</tr>
<tr>
<td>155</td>
<td>Hartmann et al., 1998</td>
<td>Tail moment</td>
<td>24</td>
<td>0.03</td>
<td>0.03</td>
<td>23.05</td>
</tr>
<tr>
<td>156</td>
<td>Šrám et al., 1998</td>
<td>Tail intensity</td>
<td>134</td>
<td>5</td>
<td>0.11</td>
<td>2.13</td>
</tr>
<tr>
<td>157</td>
<td>Moretti et al., 1996</td>
<td>Comet length</td>
<td>8</td>
<td>38.58</td>
<td>9.83</td>
<td>12.39</td>
</tr>
<tr>
<td>158</td>
<td>Van Delft et al., 2001</td>
<td>Score</td>
<td>18</td>
<td>1.3</td>
<td>0.4</td>
<td>17.4</td>
</tr>
<tr>
<td>159</td>
<td>Grover et al., 2003</td>
<td>Tail length</td>
<td>31</td>
<td>10.34</td>
<td>2.38</td>
<td>23.7</td>
</tr>
<tr>
<td>160</td>
<td>Kalaiselvi et al., 2002</td>
<td>Score</td>
<td>20</td>
<td>8.4</td>
<td>6.9</td>
<td>30.7</td>
</tr>
<tr>
<td>161</td>
<td>Garaj-Vrhovac and Kopjar, 2003</td>
<td>Tail moment</td>
<td>21</td>
<td>11.76</td>
<td>0.55</td>
<td>29.11</td>
</tr>
<tr>
<td>162</td>
<td>Biri et al., 2002</td>
<td>Score</td>
<td>9</td>
<td>113.89</td>
<td>9.75</td>
<td>109.67</td>
</tr>
<tr>
<td>163</td>
<td>Danadév et al., 2003</td>
<td>Score</td>
<td>15</td>
<td>61.27</td>
<td>14.89</td>
<td>21.57</td>
</tr>
<tr>
<td>164a</td>
<td>Başaran et al., 2003</td>
<td>Score</td>
<td>24</td>
<td>54.04</td>
<td>29.07</td>
<td>6.30</td>
</tr>
<tr>
<td>164b</td>
<td>Başaran et al., 2003</td>
<td>Score</td>
<td>14</td>
<td>63.11</td>
<td>39.5</td>
<td>8.24</td>
</tr>
<tr>
<td>165</td>
<td>Laffon et al., 2002</td>
<td>Tail length</td>
<td>11</td>
<td>42.59</td>
<td>0.96</td>
<td>14.44</td>
</tr>
<tr>
<td>166</td>
<td>Sel et al., 2003</td>
<td>Tail moment</td>
<td>26</td>
<td>1.42</td>
<td>0.32</td>
<td>17.42</td>
</tr>
<tr>
<td>167</td>
<td>Palus et al., 2003</td>
<td>Score</td>
<td>28</td>
<td>12</td>
<td>5.3</td>
<td>12.95</td>
</tr>
<tr>
<td>168</td>
<td>Maluf et al., 2001</td>
<td>Score</td>
<td>7</td>
<td>11.14</td>
<td>5.73</td>
<td>15.97</td>
</tr>
<tr>
<td>169</td>
<td>Danadév et al., 2004</td>
<td>tail length</td>
<td>57</td>
<td>10.67</td>
<td>2.62</td>
<td>45.7</td>
</tr>
<tr>
<td>170</td>
<td>Jianli et al., 2004</td>
<td>Tail moment</td>
<td>6</td>
<td>0.35</td>
<td>0.12</td>
<td>15.41</td>
</tr>
<tr>
<td>171</td>
<td>Dušinská et al., 2004</td>
<td>Score</td>
<td>7</td>
<td>105</td>
<td>87.31</td>
<td>13.57</td>
</tr>
<tr>
<td>172</td>
<td>Hoffmann and Speit, 2005</td>
<td>Tail moment</td>
<td>12</td>
<td>0.33</td>
<td>0.1</td>
<td>12.34</td>
</tr>
<tr>
<td>173</td>
<td>Touaj et al., 2002</td>
<td>Tail intensity</td>
<td>11</td>
<td>1.1</td>
<td>0.79</td>
<td>1.1</td>
</tr>
<tr>
<td>174</td>
<td>Maluf and Erdmann, 2000</td>
<td>Score</td>
<td>3</td>
<td>10.67</td>
<td>8.5</td>
<td>9.72</td>
</tr>
<tr>
<td>175</td>
<td>Poli et al., 1999</td>
<td>Comet length</td>
<td>50</td>
<td>18.36</td>
<td>16.7</td>
<td>50.48</td>
</tr>
<tr>
<td>176</td>
<td>Šárdaš et al., 2001</td>
<td>Score</td>
<td>10</td>
<td>50.5</td>
<td>21.2</td>
<td>18.40</td>
</tr>
<tr>
<td>177</td>
<td>Kopjar and Garaj-Vrhovac, 2005</td>
<td>Tail moment</td>
<td>20</td>
<td>14.19</td>
<td>0.18</td>
<td>40.13</td>
</tr>
</tbody>
</table>

*Standardized man diffrence.
| CF: confidence interval.

<table>
<thead>
<tr>
<th>Study</th>
<th>Name</th>
<th>Outcome</th>
<th>Smoker</th>
<th>Non-smoker</th>
<th>SMD* (95% CI)</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>803</td>
<td>959</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SMD = 0.71, 95% CI from 0.13 to 1.30, P = 0.02). For Type B studies (545 smokers and 656 non-smokers; median number of smokers: 15, median number of non-smokers: 16), the results are not that consistent. Whereas the fixed effects model results in the absence of any effect of smoking on DNA damage (SMD = 0, 95% CI from 0.13 to 0.12, P = 0.97), the random effects model with less weight on the large negative study, estimates the SMD as 0.47, slightly missing statistical significance (95% CI from 0.03 to 0.96, P = 0.07). However, also under the latter model, the total SMD from Type B studies is quite lower than from Type A studies.

Finally, a distinction is made between studies using image analysis or image length measurements (23 studies; Figure 5)
and those with arbitrary determinations of a score (15 studies; Figure 6). In the first 23 studies (576 smokers and 700 non-smokers; median number of smokers and non-smokers: 20), again the overall results are not quite consistent across the models, at least partly caused by the study by Šrám et al. (156). Under the fixed effects model, the total SMD = 0.02 (95% CI from −0.10 to 0.14, P = 0.72), and the random effects model produces an estimate equal to 0.42 (95% CI from −0.12 to 0.97, P = 0.13). However, whereas 23 studies with quantitative outcome did not reach a formal proof of increased DNA damage in smokers, but only indicate a tendency, the pooled results of 15 studies using a scored outcome (227 smokers and 259 non-smokers; median number of smokers and non-smokers: 15) are doubtlessly in favor of a smoking effect. Computations under the fixed effects assumptions yield SMD = 0.72 (95% CI from 0.50 to 0.94, P < 0.0001), under the random effects assumptions SMD = 0.80 (95% CI from 0.37 to 1.24, P = 0.0003).

Table I summarizes 13 studies which could not be included into our meta-analysis because information necessary for data transformation were missing. However, these studies indicated the result of a statistical test and presented a P-value. Based on this information, seven studies reported a genotoxic effect in blood cells of smokers. One of these statements was based on a higher number of ‘undamaged’ cells in non-smokers (131). One study revealed an equivocal result, i.e. an effect in smokers was determined with one statistical analysis (‘bivariate correlation’) but not with another (‘multiple linear regression’) (124). Both types of analysis are not very common in such kind of studies. Three out of these thirteen studies reported no significant difference between smokers and non-smokers and two studies found significantly higher levels of DNA damage in non-smokers (132,134).

Discussion

Our meta-analysis suggests that smoking has an effect on DNA migration in the comet assay. We only considered studies investigating peripheral blood because this approach is mainly used in human monitoring. We did not differentiate between studies using whole blood (measuring effects in leukocytes) or isolated lymphocytes. Although both methods basically seem to be appropriate, the use of whole blood is recommended when exposure conditions are at steady state (8). Furthermore, the procedure of lymphocyte isolation may induce mechanical or oxidative damage in the cells, thus increasing the background level of damage and the variability of the test. We only considered the standard procedure of the alkaline comet assay, which measures ‘strand breaks’ (i.e. DNA single- and double-strand breaks, ALS and repair incisions). It might be that the inclusion of lesion-specific enzymes for the demonstration of (oxidative) DNA base damage can render the test more sensitive and theoretically lead to a better detection of smoking-related genotoxic effects. However, this modification of the comet assay increases assay variability and the few studies using this approach also revealed conflicting results (143,172,180,181).

Although the comet assay has been shown to be a sensitive in vitro and in vivo genotoxicity test, its sensitivity in human biomonitoring is frequently stated but is not proven and has to be questioned. The comet assay just measures strand breaks and ALS, i.e. primary DNA lesions which are short-lived and rapidly repaired in mammalian cells. Bulky DNA adducts which are formed from many mutagens/carcinogens (and which are also contained in tobacco smoke), do not directly cause comet assay effects but only increase DNA migration.
when they persist and are repaired by excision repair in cells under investigation (182). Therefore, the comet assay might be less suited for the detection of low levels of chronic exposure towards genotoxins and it is understandable that smokers do not exhibit clear effects in the comet assay.

The reasons for the conflicting results obtained with the comet assay for the effect of smoking have been discussed in recently published reviews (6,7). One important aspect is the statistical power of the studies. Low statistical power can be the reason for the failure of some studies to demonstrate an effect, and a meta-analysis can help to circumvent this by pooling the available studies. However, when a statistically significant effect is stated, most of the reported positive effects are small, and the biological relevance of the effect is usually not discussed. Given the high assay variability, an increase by 10% from the control value (141) may be statistically significant but has no meaning at all considering the biology of the test. Furthermore, this latter issue, of course, cannot be solved by a meta-analysis. Clear differences between smokers and non-smokers were also reported in small studies (144,150,154,174). This may be partly due to chance, as in small studies effect estimates have lower precision and may vary substantially between studies. Furthermore, considering the whole database, it has to be questioned that effects of this magnitude are actually related to smoking as long as there are no indications that the cigarettes smoked in these studies do contain relevant amounts of strong genotoxins which are not present in other cigarettes. However, no data are available of a possible impact of the brand and composition of cigarettes on their potential genotoxic effects. Furthermore, the available studies did not indicate a relationship between the number of cigarettes smoked and the effect in the comet assay. Whereas one group reported significant increase in DNA damage in smokers smoking one or more cigarettes (152), other studies did not find differences between non-smokers and smokers smoking >10 cigarettes/day (168,173). We recently reported a negative result in a well-controlled study with heavy smokers (>20 cigarettes/day) (172). Because the information is not generally provided, the amount of cigarette smoking was not particularly considered in our meta-analysis.

It has been observed that the positive studies seem to cluster in the southern part of Europe (6) and led to the speculation that geographical variation in smoking habits may be relevant. It is also likely that lifestyle factors, diet (e.g. antioxidants) and genetic differences (e.g. genetic polymorphisms in genes involved in metabolism of xenobiotics and DNA repair) influence the effects in the comet assay. However, the limited data available point to the complexity of this matter and also show conflicting results (143,158,172,180,181).

In our opinion, the most critical aspect with regard to the evaluation of comet assay results from human monitoring studies is the assay variability and the poor quality of many studies. Many studies do not clearly state that the evaluation is based on coded slides and that slides from smokers and non-smokers are processed and evaluated concurrently. An inappropriate study design (e.g. no masking of the evaluation) opens the possibility that the results are influenced by secondary (e.g. psychological) factors. Also minor and trivial changes in the comet assay procedure (e.g. the exhaustion of the fluorescent lamp) may lead to shift in the sensitivity with time. It has to be clearly stated that the quality of the study design and the performance of the study, regardless whether the study was positive or negative, could not be considered in this meta-analysis because this information is included only in a minority of studies. Measures of quality assurance as suggested by the IPCS (8) are hardly ever used. Given the wide inter-individual range of baseline DNA-damage levels within any group, as well as intra-individual variability resulting from endogenous and exogenous factors, quality assurance to demonstrate the adequacy of the methodology used in the study and reproducibility of observed effects is indispensable.

We had to exclude 19 studies from our analysis because the lack of raw data or the lack of basic descriptive statistical measures did not allow a formal joint evaluation under standard models. The bias caused by omitting these studies is difficult to be assessed. However, among the excluded 14 studies presented in Table I, 7 reported increased DNA damage in smokers, 4 did not find any effect and 2 measured increased levels of DNA damage in non-smokers. One study (124) revealed an equivocal result depending on the statistical method. Furthermore, five studies were excluded (136–140) because they only mentioned the effect of smoking without showing data. Out of these five studies, one reported a genotoxic effect in smokers (139), whereas the other four did not find a difference between smokers and non-smokers. Taken together, among these 19 studies which were excluded from our meta-analysis because they did not provide necessary information, 8 found an effect in smokers, 8 found no difference, 2 found higher effects in non-smokers and 1 reported an equivocal result. Thus, this additional information indicates that the formal exclusion of studies did not fundamentally change the ratio of available positive and negative studies which investigated genotoxic effects of cigarette smoking in peripheral blood cells with the comet assay.

A subdivision of the publications lead to two interesting findings:

(i) the effect of smoking was more pronounced when smoking was investigated as a genotoxic exposure (Type A studies) and there was at most a ‘borderline’ smoking-associated effect when smoking was investigated as a potential confounding factor in studies with occupational exposure towards genotoxins (Type B studies). This difference is surprising because there seems to be no fundamental difference between these two study types with respect to the size of the samples or and the amount of exposure. The argument could be raised that Type A studies, in contrast to Type B studies, have been planned in order to detect effects of certain magnitudes and thus might have higher power than Type B studies, where, e.g. uncontrolled factors may increase within-group variability. In our meta-analysis, we have observed that Type A studies indeed comprise slightly more smokers and non-smokers than Type B studies. However, in our meta-analysis, we have more than twice as many probands in Type B studies than in Type A studies. This is supposed to compensate for a smaller average power of the Type B studies. Indeed, the length of the confidence intervals for SMD is about the same in both study groups. We see that effect sizes in Type B studies are on average less pronounced than in Type A studies. Besides that, no study considered here presented a formal a priori sample size determination. Although the effects of smoking should not influence the publication of occupational studies, it might be that there is a publication bias in favor of positive effects for studies investigating genotoxic effects of smoking.
(ii) An effect of smoking could not be formally demonstrated (at most a tendency could be detected) when the evaluation was based on image analysis and direct length measurements, but studies using a visual score clearly indicated an effect after pooling. This difference is unexpected because evaluation by visual scoring (arbitrary assignment to damage classes) should not be more sensitive than image analysis or direct length measurements. A direct comparison of slides evaluated by image analysis and visual scoring indicated a close correspondence between the two approaches but a little lower precision of visual scoring (183). This means that for the same level of reliability the number of subjects should be higher in studies relying on visual scoring alone. However, this is not the case here (in median, 15 probands/group in contrast to 20 probands/group with image analysis) and positive results were also reported in studies using visual scoring and a small number of subjects (144,174).

Taken together, our meta-analysis indicates that on the basis of 38 studies (from 37 publications) smoking has a DNA-damaging effect on peripheral blood cells as measured by the comet assay. However, this effect is more pronounced when smoking was investigated as a genotoxic exposure and only assessing a genotoxic exposure and only investigating the damaging effect on peripheral blood cells as measured by the comet assay. 

Acknowledgement
We wish to thank Mrs Yvonne Spaziano for valuable help with the data analysis.

References


in pesticide production utilizing the comet assay. Mutagenesis, 18, 201–205.


Received on August 10, 2005; revised on October 17, 2005; accepted on October 17, 2005.