Low integrated DNA repair score and lung cancer risk

Ziv Sevilya¹, Yael Leitner-Dagan¹, Mila Pinchev², Ran Kremer³, Dalia Elinger¹, Hedy S. Rennert², Edna Schechtman⁴, Laurence Freedman⁵, Gad Rennert², Tamar Paz-Elizur*¹ and Zvi Livneh*¹

¹Dept of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel; ²Dept of Community Medicine and Epidemiology, Carmel Medical Center, Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, and Clalit Health Services National Cancer Control Center, Haifa 34362, Israel. ³Dept of General Thoracic Surgery, Rambam Health Care Campus, Haifa 31096, Israel; ⁴Dept. of Industrial Engineering and Management, Ben Gurion University of the Negev, Beer Sheva 84105, Israel; ⁵Biostatistics Unit, Gertner Institute for Epidemiology and Public Health Policy Sheba Medical Center Tel Hashomer 52621, Israel.

Running title: Low DNA repair score and lung cancer risk

Keywords: DNA repair; risk biomarker; risk factor; lung cancer; cancer prevention; CT screening; low-dose CT;

Word count, number of figures and tables: 4920 words, 1 Figures and 4 Tables.

Type of article: Research Article

*Corresponding authors: zvi.livneh@weizmann.ac.il (Zvi Livneh), tamar.paz- elizur@weizmann.ac.il (Tamar Paz-Elizur). Dept. of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel. Tel: +972-8-934-3203; Fax: +972-8-934-4169.
Financial support: NIH/NCI/EDRN (#1 U01 CA111219) to Z. Livneh, T. Paz-Elizur and G. Rennert, and the Flight Attendant Medical Research Institute, Florida to Z. Livneh and T. Paz-Elizur.

Disclosure of Potential Conflict of interest: The corresponding authors (Z. Livneh and T. Paz-Elizur) have a patent on the OGG1 risk factor for lung cancer, and a patent application pending for the OGG1, MPG and APE1 panel of DNA repair biomarkers.
Abstract

DNA repair is a prime mechanism for preventing DNA damage, mutation and cancers. Adopting a functional approach we examined the association with lung cancer risk of an integrated DNA repair score, measured by a panel of three enzymatic DNA repair activities in peripheral blood mononuclear cells. The panel included assays for AP endonuclease 1 (APE1), 8-oxoguanine DNA glycosylase (OGG1) and methylpurine DNA glycosylase (MPG), all of which repair oxidative DNA damage as part of the base excision repair pathways. A blinded population-based case-control study was conducted with 96 lung cancer patients and 96 control subjects matched by gender, age (±1 year), place of residence, and ethnic group (Jews/Arabs). The three DNA repair activities were measured, and an integrated DNA repair OMA (OGG1, MPG & APE1) score was calculated for each individual. Conditional logistic regression analysis revealed that individuals in the lowest tertile of the integrated DNA repair OMA score had an increased risk of lung cancer compared to the highest tertile, with OR 9.7; 95% CI, 3.1-29.8; \( P<0.001 \), or OR 5.6; 95% CI, 2.1-15.1; \( P<0.001 \) after cross validation. These results suggest that pending validation, this DNA repair panel of risk factors may be useful for lung cancer risk assessment, assisting prevention and referral to early detection by technologies such as low-dose CT scanning.
Introduction

Prevention is arguably the most effective way to overcome disease and improve human health. This has been demonstrated by the dramatic decrease in heart disease mortality over the last 20 years, a significant part of which was due to the identification of mechanism-based risk biomarkers, and taking effective measures to treat them (1). Despite the advance in the molecular mechanisms that underlie cancer development, progress in cancer prevention has been slow (2). A key reason is the lack of effective mechanism-based host risk factors (e.g., risk biomarkers) that estimate to one’s risk to develop cancer (3). DNA repair plays a critical role in preventing cancer via its DNA safeguarding activities, which prevent mutations and genomic instability (4) that are critical for cancer (5-8). This suggests that DNA repair is a promising target for mechanism-based cancer prevention approaches (9-12).

Aiming at harnessing DNA repair for cancer prevention, it was hypothesized that variations exist among individuals in their ability to repair DNA, and that those variations translate into differences in cancer predisposition, such that low effectiveness in DNA repair capacity will cause higher cancer risk. To measure such inter-individual variations in DNA repair we (13-18), and others (19-21), took the approach of measuring the enzymatic activity of specific DNA repair enzymes, using PBMC as the reporter tissue. Our primary focus was on DNA repair enzymes that target oxidative DNA damage, which has been implicated as a major factor in carcinogenesis (22). This approach assumes that (a) DNA repair activity in PBMC is a legitimate surrogate for DNA repair in other tissues, and (b) each individual has a personal baseline capacity to repair DNA. These assumptions are supported by previous results showing that the activity of the DNA repair enzyme OGG1 in PBMC correlates to the activity in lung tissue (13), and that individuals exhibit a personal OGG1 activity that is stable over several years (13) and
does not significantly change years after recovery from cancer (15). Using this approach we found that low OGG1 enzymatic activity is associated with the risk of lung cancer (13), and head and neck cancer (15). In addition, we found that increased enzymatic activity of MPG, a DNA repair enzyme of the DNA glycosylases family, which removes from DNA multiple alkylation and oxidative lesions, is associated with increased lung cancer risk (18). This surprising finding illustrates the notion that imbalances in DNA repair (23-25), and not necessarily low activity of DNA repair, can be involved in cancer risk.

Here we present evidence for a third enzymatic DNA repair risk factor for lung cancer, namely low enzyme activity of AP endonuclease 1 (APE1). APE1 is of special interest because of its central role in base excision repair, a pathway to which OGG1 and MPG also belong (26-29); Fig. 1A). Moreover, we show that a personalized integrated DNA repair score, composed of a combination of the new APE1 risk factor and the OGG1 and MPG risk factors, is strongly associated with lung cancer risk, and pending validation may provide a tool for improved selection criteria of extra-high risk individuals for early detection approaches such as low-dose CT scanning (30, 31).

Materials and Methods

Study participants

The study participants have been previously described (18). They included 100 patients recently diagnosed with lung cancer (case patients) recruited from the Rambam Medical Center in Haifa, located in northern Israel, and 100 subjects with no diagnosis of lung cancer (control subjects) who were randomly selected enrollees of Clalit Health Services (CHS) identified from the same geographical area. The controls were individually matched to the cases by gender, year
of birth, place of residence (defined by primary clinic location), and ethnic group (Jews vs. non-
Jews). Controls were excluded only if they had a former diagnosis of lung cancer. Participants
provided written informed consent at the time of recruitment, and were interviewed in person to
obtain information about their personal and family history of cancer, and smoking history.
Diagnoses of lung cancer were made independently by the diagnosing hospitals and included
information on histological type, TNM staging and tumor grade. The institutional review board
at Carmel Medical Center, Haifa, approved all procedures. Blood samples were drawn from
cases prior to the operative procedure or any treatment intervention. The case patients and
control subjects had mean ages of 67 years, included 60% males and 40% females, and were
77% Jews and 23% non-Jews. The controls comprised 22%, 28% and 50% current, former and
never smokers, respectively and the cases 37%, 36% and 24%, respectively (18). There was no
matching for smoking status, but adjustment was performed for smoking in the statistical
analyses.

**Isolation of peripheral blood mononuclear cells, and preparation of protein extracts**

Each study participant provided a sample (17mL) of peripheral blood collected in two
ACD blood collection tubes (BD vacutainers catalog#364606, NJ USA). Blood samples were
kept at room temperature, and processed 18-24 hours after collection to isolate PBMC,
especially as previously described (18). Protein extracts were prepared from PBMC as
previously described (18).
Assays of enzymatic DNA repair activity

The fluorescence-based DNA repair assays were done with DNA substrates tagged with a 3’-yakima yellow fluorescent tag, each containing the appropriate site-specific lesion: a synthetic abasic site (also called apurinic/apyrimidinic site; AP site) for APE1, an hypoxanthine (Hx) for MPG, and 8-oxoguanine (8-oxoG) for OGG1. The assays monitor the specific incision at the lesion, followed by capillary gel electrophoresis analysis using an AB3130XL genetic analyzer. The assays were done in a robotic platform, in which liquid handling of the nicking reactions were performed automatically by a Freedom EVO 200 robot (Tecan) and Freedom EVOware software (Tecan).

APE1 assay. The assay used for measuring APE1 enzymatic activity is in principle similar to previously published assays (32, 33). Its development will be described elsewhere, but the final reaction conditions were as follows: The DNA substrate used was 5’- G GTG CAT GAC ACT GTF ACC TAT CCT CAG CG Y -3’ (F = furanyl abasic site; Y = Yakima yellow tag) annealed to 5’- CG CTG AGG ATA GGT TAC AGT GTC ATG CAC C -3’. The reaction mixture (20 μl) contained 75 mM Tris (pH 7.8), 0.1 mM EDTA, 9 mM MgCl2, 42.4 mM KCl, 0.25 μg/μl bovine γ-globulin, 0.25% PVA, 0.25 mM Spermidine; 0.05 mM Spermine; 35 nM substrate and 0.015 ng/μl protein extract. The reaction was carried out at 37 °C for 15 minutes, after which it was stopped by heat inactivation at 95°C for 2 minutes. One unit of APE1 activity is defined to cleave 1 fmol of DNA substrate in 1 hour at 37 °C, under these conditions.

MPG assay. The DNA substrate used for the MPG assay was 5’- GT CCG GTG CAT GAC ACT GTX ACC TAT CCT CAG CG Y -3’ (X = hypoxanthine; Y = Yakima yellow tag) annealed to 5’- CG CTG AGG ATA GGT TAC AGT GTC ATG CAC CGG AC -3’. The standard reaction mixture (20 μl) contained 50 mM MOPS (pH 6.8), 30 mM Tris (pH 7.1), 2 mM...
EDTA, 0.5 mM MgCl₂, 36 mM KCl, 1 μg/μl bovine γ-globulin, 0.3% PVA, 15 nM substrate and 45 ng/μl protein extract. The reaction was carried out at 37 °C for 15 minutes, after which it was stopped by heat inactivation at 95°C for 2 minutes. The reactions were treated with 100 mM NaOH for 30 minutes at 37 °C. One unit of MPG activity is defined to cleave 1 fmol of DNA substrate in 1 hour at 37°C, under these conditions.

**OGG1 assay.** The DNA substrate used for the OGG1 assay was 5’-CCG GTG CAT GAC ACT GTZ ACC TAT CCT CAG CG Y -3’ (Z = 8-oxoG; Y = Yakima yellow tag) annealed to 5’-CGC TGA GGA TAG GT CACA GTG TCA TGC ACC GG -3’. The standard reaction mixture (10μl) contained 50mM Tris (pH 7.1), 1mM EDTA, 115mM KCl, 1μg/μl bovine γ-globulin, 100nM PolydA-polydT, 12.5nM substrate and 0.2-0.5mg/ml protein extract. The reaction was carried out at 37°C for 30 minutes, after which it was stopped by heat inactivation at 95°C for 2 minutes. The reactions were treated with 100mM NaOH for 30 minutes at 37°C. One unit of OGG1 activity is defined to cleave 1fmol of DNA substrate in 1 hour at 37°C, under these reaction conditions.

All enzyme activities measured in the case-control specimens were assayed also in two standard extracts that were used in each round of the assays, and served as reference normalizing extracts. These standard extracts were prepared from blood bags obtained from the blood bank.

**Measurement of baseline APE1, OGG1 and MPG enzyme activities.** Baseline enzymatic activities of APE1, OGG1 and MPG were measured in the same PBMC specimens, obtained from 10 blood bank bags, using the assays described above. The results obtained for the average activities of OGG1 and MPG were 5.8 (range: 5.0-6.7) and 167 (range: 146-189) units/μg protein, respectively, whereas surprisingly, APE1 activity was 1,121,000 units/μg protein (range: 994,000-1,247,000), 5 orders of magnitude higher than OGG1 activity (Fig. 1B).
Statistical analysis

The odds ratio of lung cancer was estimated for APE1 tests using conditional logistic regression models adjusting for smoking status with or without OGG1 and MPG activities. This method takes account of the matched design of the study and therefore automatically adjusts also for the matching variables, sex, age, place of residence and ethnicity. Two models were used to estimate the odds ratio. In the first model the explanatory variable was treated as a continuous variable (assuming a linear relation with the log odds), and in the second model the variable was categorized into three groups according to the tertiles of the controls. In the latter case a test for a linear trend across tertiles in the log odds was conducted using scores of 1, 2 and 3 for the three tertile groups.

The odds ratio of lung cancer was estimated for the OMA score using conditional logistic regression models adjusting for smoking status. Odds ratios were estimated for the continuous variable (assuming a linear relation with the log odds), and categorized into three groups, as with APE1. Because the weights for each component used in the OMA score were chosen to optimize strength of association of the score with lung cancer for the observed data, a cross-validation procedure was applied to remove the resulting bias in the odds ratios. A ‘cross-validated’ OMA score was obtained for each individual by omitting the individual and his/her matched pair and running the conditional logistic regression model on the remaining pairs so as to obtain new weights for the OMA score. The cross-validated OMA score for the omitted pair was then calculated using these new weights. We did this for each matched pair, and then ran the conditional logistic regression on the set of cross-validated OMA scores instead of the original scores. All the statistical analyses were performed using S-Plus (TIBCO Software Inc.) and /or SAS software (version 9.2; SAS Institute Inc.).
Results

APE1 activity characteristics in lung cancer case patients and matched control subjects

We first concentrated on examining whether inter-individual variations in APE1 activity, like OGG1 and MPG, are also associated with lung cancer risk, thereby increasing the number of risk biomarkers to three, and likely provide a better risk estimate. The results of measuring APE1 enzyme activity of the case-control lung cancer patients and matched control subjects are presented in Table 1 and Supplementary Fig. 1A. The inter-individual variation among the controls was 4.9-fold, ranging from 435 to 2137 units/ng protein, and among case patients 3.3-fold, ranging from 432 to 1410 units/ng protein. The mean APE1 specific activity in case patients was 787 (95% CI 752-823) units/ng protein, significantly lower than the 896 (95% CI 849-943) units/ng protein mean activity in control subjects (P=0.0007; Table 1). APE1 activity in the controls appeared lower in males (859 units/ng protein) than in females (951 units/ng protein), but not among cases (783 units/ng protein in males versus 794 in females)(test for gender-disease interaction: P=0.13; Table 1). In addition, APE1 activity was lower in those aged >65y than in those aged ≤65y (P=0.018) (Table 1). There was no difference in APE1 activity between adenocarcinoma and squamous cell carcinoma (Table 1). To determine whether smoking status is associated with APE1 activity level, the mean APE1 activity level was calculated for each smoking status in case patients and control subjects. As can be seen in Table 1, the mean levels of APE1 activity among current, past and never smokers were similar, suggesting that smoking did not significantly affect the level of APE1 enzyme activity in PBMC. In addition, no significant interaction between smoking status and APE1 activity was found when a test for interaction was performed (Table 1).
**Decreased APE1 enzymatic activity is associated with increased risk of lung cancer**

The association between the level of activity of APE1 and the probability of having lung cancer, adjusted for smoking status, was further explored using conditional logistic regression. When APE1 activity was used as a continuous variable the adjusted odds ratio for lung cancer associated with 1 standard deviation (1 SD; 239 units)\(^1\) decrease in APE1 activity was 2.0 (95% CI, 1.3-3.1; \(P=0.002\)) (Table 2). When APE1 activity values were divided into tertiles using the controls’ values, the adjusted odds ratio for the lowest tertile versus the highest was 3.7 (95% CI, 1.5-8.8; \(P=0.003\)). A test for trend over the three tertiles was also significant (\(P=0.003\), Table 2). The significant associations obtained with the two different models strengthen the conclusion that low APE1 enzymatic activity is associated with lung cancer risk.

**The integrated DNA repair OMA score is strongly associated with lung cancer risk**

When measuring the three DNA repair activities in an individual, it is useful to form an integrated DNA repair score of the three with appropriate weights, because this takes into account situations in which only one or two of the DNA repair activities are in the range that is associated with higher cancer risk. The weights of the score were obtained directly from the estimated log odds ratios in the fit of the conditional logistic regression model including OGG1, MPG and APE1 (with smoking status included in the model; Supplementary Table 1), and the score was defined as \(0.00425\times\text{APE1} + 0.5419\times\text{OGG1} - 0.02541\times\text{MPG}\). We term this the OMA score (for OGG1, MPG and APE1), and calculated its value for each study participant. The distributions of the OMA score among study participants showed a clear shift of case patients to lower OMA score values compared to control subjects (Supplementary Fig. 1B). The OMA

\(^1\) Expressing the variable in standard deviation (SD) units (standardized coefficients) allows meaningful comparisons to other DNA repair assays.
scores ranges were 1.40-10.42 and 0.53-7.01 units in controls subjects and case patients, respectively. The mean OMA score in patients was 2.8 (95% CI, 2.6-3.0), significantly lower than the value of 3.6 (95% CI, 3.4-3.8) in control subjects ($P<0.0001$; Table 3). The integrated DNA repair OMA score in adenocarcinoma cases was 2.9 (95% CI, 2.7-3.2), higher than in SCC cases, where it was 2.5 (95% CI, 2.2-2.7; $P=0.04$). There was no appreciable difference in the OMA score between males and females, and between subjects $>65$ and $\leq 65$ years old. In addition, there was no appreciable difference between never, former and current smokers, and no interaction between the OMA score and smoking status.

The association between the OMA score and the probability of having the disease, adjusted for smoking status, was calculated using conditional logistic regression with the OMA score as a continuous variable. As can be seen in Table 4, the adjusted odds ratio for lung cancer associated with a 1.18 unit decrease (1 standard deviation) in OMA score was significantly greater than 1 (OR 3.2; 95% CI, 1.7-5.7; $P<0.001$) (Table 4). We used cross-validation to adjust the estimated odds ratios so that they will reflect the likely performance of the score in a new dataset and obtained odds ratio of 2.7 (95% CI, 1.6-4.5; $P<0.001$), which is smaller, but still very strong (Table 2). The strength of the association of reduced OMA score with lung cancer risk is further illustrated in a logistic regression model using tertiles of OMA score among control subjects. After adjusting for smoking status, the odds ratio for lung cancer of individuals in the lowest tertile (versus the highest tertile) of the DNA repair OMA score was 9.7 (95% CI, 3.1-29.8; $P<0.001$), and after cross validation OR 5.6 (95% CI, 2.1-15.1; $P<0.001$), compared to individuals in the upper tertile (Table 4).
Discussion

Because the important role of oxidative stress in carcinogenesis, imbalances in DNA repair enzymes that repair oxidative DNA damage are attractive candidate risk factors for lung cancer. The diversity of oxidative DNA damage suggests that a single DNA repair risk factor will not provide a sufficiently strong risk prediction, and rather a panel of risk factors will be needed. We have thus chosen APE1 enzymatic activity, in addition to the OGG1 and MPG activities that we have previously studied, because of its key function in BER (4, 27-29, 34), and its involvement in single-strand break repair (35, 36) and nucleotide incision repair (NIR) (37). In addition APE1 is a redox factor via its REF domain (38).

Somewhat surprisingly, APE1 enzyme activity in PBMC was 4-5 orders of magnitude higher than the activities of OGG1 and MPG (Fig. 1B), likely because its involvement in multiple DNA repair pathways as described above. These high APE1 activities are not saturating, and the 4.9-fold range of APE1 activity that we have observed amongst control subjects appears to have biological significance, as indicated by the significant association of low APE1 activity and lung cancer risk. This association provides support for the importance of oxidative DNA damage in the development of lung cancer, and is intuitively understandable since low repair causes higher mutation rates due to unrepaired oxidative DNA damage, thereby facilitating carcinogenesis. A role of low APE1 in facilitating carcinogenesis is at first glance inconsistent with the observations that APE1 is overexpressed in several types of cancers, also associated with resistance to chemotherapy (38-41). To explain this we hypothesize a dual role for DNA repair in carcinogenesis. In normal cells low APE1 activity facilitates carcinogenesis by increasing mutations and genomic instability, however, as cells progress along the carcinogenesis path, which involves robust proliferation, cells with up-regulated DNA repair
may be selected because of the advantage that it provides during replication.

The OMA score presented in this study enables to measure one’s ability to repair oxidative DNA damage, representing a personalized DNA repair capacity. The strong association of low OMA score with lung cancer risk is highlighted in Supplementary Fig. 1C that shows the odds ratio calculated as a function of the percentiles of the distribution of OMA score among the control subjects. The odds ratio for a person at the 10th (lowest) percentile versus a person at the 90th percentile of the distribution was 11.5 (95% CI, 3.6-36.3), or 7.6 (95% CI, 2.7-20.9) after adjustment by cross validation. The strong association of the integrated DNA repair score suggests that it can potentially be used for lung cancer risk assessment, thereby aiding lung cancer prevention. Importantly, the fact that low integrated DNA repair OMA score is a strong risk factor, which is independent of smoking status, suggests that it may serve as a powerful component, on top of smoking status, in evaluating one’s risk to develop lung cancer. This can be used to estimate personalized risk of a smoker to develop lung cancer, and provide motivation for smoking cessation and other risk reduction activities.

Like in any epidemiological study, validation of the results is needed in independent studies. In this context, it should be pointed out that the present OGG1 results validate a previous study that we have performed on different groups of lung cancer patients and healthy control subjects (13). A limitation of the present study is its relatively small sample size, which limits the ability to perform sub-group analysis. In addition, a prospective study is needed for definitively proving that low OMA DNA repair score has predictive power for lung cancer.

It was recently reported that low-dose computed tomography (LD-CT) screening for early detection of lung cancer enabled identification and treatment of very early lung cancer, leading to a 20% decrease in lung cancer mortality. However, the low rate of 4% true positives in
this detection method complicates its broad usage for screening purposes (30) (NLST study; see also the I-ELCAP study (31)). A useful paradigm would be to use the integrated DNA repair OMA score as a mean to improve selection criteria for LD-CT screening for early detection of lung cancer, by identifying individuals at extra-high risk. Pending validation, such a strategy is expected to significantly increase the rate of true-positives, thereby avoiding unnecessary invasive procedures and improving cost-benefit (42-44). In addition, having the OMA personalized measure for DNA repair activity raises the possibility of searching for drugs that will improve the effectiveness of DNA repair, potentially providing protection against lung cancer.

In summary, our study identified a panel of DNA repair risk factors, which when combined produced an individual risk score that is strongly associated with lung cancer risk. Pending validation it can be used to motivate or select individuals for intensive primary or secondary prevention intervention.

Acknowledgements

This work was funded by grants from NIH/NCI/EDRN (#1 U01 CA111219) to ZL, TPE and GR, and from the Flight Attendant Medical Research Institute, Florida to ZL and TPE.
References


carcinogen-induced DNA adducts and lung cancer risk: a molecular epidemiologic study. J Natl

capacity of lymphoblastoid cell lines derived from sisters discordant for breast cancer. J Natl

13. Paz-Elizur T, Krupsky M, Blumenstein S, Elinger D, Schechtman E, Livneh Z. Reduced


repair of the oxidative 8-oxoguanine DNA damage and the risk of head and neck cancer. Cancer

Development of an enzymatic DNA repair assay for molecular epidemiology studies:
Distribution of OGG activity in healthy individuals. DNA Repair. 2007;6:45-60.

17. Paz-Elizur T, Sevilya Z, Leitner-Dagan Y, Elinger D, Roisman LC, Livneh Z. DNA repair of
oxidative damage in human carcinogenesis: potential application for cancer risk assessment and

Methylpurine DNA Glycosylase and OGG1 DNA Repair Activities: Opposite Associations With


Figure Legends

Figure 1. Base excision repair and enzymes studied in the present work. A, Each of the DNA glycosylases OGG1 and MPG initiates a sub-pathway of BER by removing from DNA its damaged base target (the black circle and triangle). The abasic site thus formed is incised by the AP endonuclease APE1, followed by further processing, repair synthesis and ligation to complete the repair. APE1 also initiates the repair of primary abasic sites formed in DNA by spontaneous or oxidation-induced base loss. B, Mean DNA repair enzyme activity in PBMC specimens obtained from 10 control subjects, which were assayed for OGG1 acting on 8-oxoguanine, MPG acting on hypoxanthine, and APE1 acting on a furanyl abasic site. The numbers inside the columns are the values of the mean enzymatic activity in units/μg protein, whereas the numbers above the columns represent the activities relative to OGG1 activity. Assay conditions are described under Materials and Methods.
Table 1. Distribution of selected characteristics and APE1 activity value in lung cancer patients and control subjects*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control subjects (n=100)</th>
<th>Case patients (n=100)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>APE1 mean (95% CI)</td>
<td>APE1 mean (95% CI)</td>
<td></td>
</tr>
<tr>
<td>All~</td>
<td>99</td>
<td>99</td>
<td>0.0007†</td>
</tr>
<tr>
<td>SQCC</td>
<td>30</td>
<td>747 (686-807)</td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>45</td>
<td>805 (757-854)</td>
<td></td>
</tr>
<tr>
<td>P interaction#</td>
<td></td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤65</td>
<td>40</td>
<td>933 (867-1000)</td>
<td>0.018</td>
</tr>
<tr>
<td>&gt;65</td>
<td>59</td>
<td>871 (807-936)</td>
<td></td>
</tr>
<tr>
<td>P interaction#</td>
<td></td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>59</td>
<td>859 (811-907)</td>
<td>0.14</td>
</tr>
<tr>
<td>Female</td>
<td>40</td>
<td>951 (860-1042)</td>
<td></td>
</tr>
<tr>
<td>P interaction#</td>
<td></td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td>0.40</td>
</tr>
<tr>
<td>Never smoked</td>
<td>50</td>
<td>927 (851-1002)</td>
<td></td>
</tr>
<tr>
<td>Past smoker</td>
<td>27</td>
<td>867 (792-942)</td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>22</td>
<td>863 (780-945)</td>
<td></td>
</tr>
<tr>
<td>P comparing subgroups ^</td>
<td></td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>P interaction#</td>
<td></td>
<td>0.55</td>
<td></td>
</tr>
</tbody>
</table>

* APE1 activity was measured as described under “Materials and Methods”. One participant did not have a known APE1 value. This participant and the matched control were excluded from the analysis. Three case participants did not have a known smoking status.

~ Of the 100 lung cancer cases, 30 had squamous cell carcinoma (SQCC), 46 had adenocarcinoma, 14 BAC, 4 adenosquamous carcinoma, 4 adenoBAC; 1 small cell carcinoma, and 1 unknown histology.

† Analysis of covariance comparing cases with controls, with matched pair and smoking status as a covariate.

§ Analysis of covariance comparing histological type within cases, with smoking status, age (continuous) or gender as covariates.

# Test for interaction between case-control status and the variable of interest.

^ Analysis of covariance comparing subsets defined by the variable of interest and stratified by cases and controls, with smoking status, age (continuous) or gender as covariates, as appropriate.
**Table 2.** Conditional logistic regression analysis of APE1 activity value in lung cancer patients and control subjects*

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. of control subjects (%)</th>
<th>No. of case patients (%)</th>
<th>Adjusted† OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>APE1 (per 239 U decrease)‡§</td>
<td>96 (100.0)</td>
<td>96 (100.0)</td>
<td>2.0 (1.3-3.1)</td>
</tr>
<tr>
<td>APE1 (by tertiles)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;961 U</td>
<td>32 (33.3)</td>
<td>14 (14.6)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td>795–961 U</td>
<td>32 (33.3)</td>
<td>29 (30.2)</td>
<td>1.7 (0.7-3.9) *P=0.22</td>
</tr>
<tr>
<td>≤795 U</td>
<td>32 (33.3)</td>
<td>53 (55.2)</td>
<td>3.7 (1.5-8.8) *P=0.003</td>
</tr>
</tbody>
</table>

*OR = odds ratio; CI = confidence interval.
†Conditional logistic regression for matched sets adjusted for smoking status (smoker, ex-smoker, never smoker). Three case patients did not have known smoking status and one control subject did not have known APE1 value. These participants along with their matched subjects were excluded from the analysis.
‡APE1 activity was measured as described in the “Materials and Methods” section and was first fitted in the conditional logistic regression model as a continuous variable and with adjustment for smoking status. The odds ratio for smoking that was obtained with this model, was: ex-smoker v never smoker: 3.2 (95% CI= 1.3-7.6); current smoker v never smoker: 3.0 (1.3-6.8).
§ 239 U represents 1 standard deviation in the control group.
|| Tertiles of the control subjects’ values. The upper tertile was chosen as the referent.
**Table 3.** Distribution of selected characteristics and combined OMA integrated DNA repair score in lung cancer patients and control subjects*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control subjects (n=100)</th>
<th>Case patients (n=100)</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OMA score mean (95% CI)</td>
<td>OMA score mean (95% CI)</td>
<td></td>
</tr>
<tr>
<td>All~</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>99</td>
<td>99</td>
<td>&lt;0.0001†</td>
</tr>
<tr>
<td>SQCC</td>
<td>30</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>2.5 (2.2-2.7)</td>
<td>2.9 (2.7-3.2)</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.35</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤65</td>
<td>40</td>
<td>40</td>
<td>0.67</td>
</tr>
<tr>
<td>&gt;65</td>
<td>59</td>
<td>59</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>3.6 (3.3-3.9)</td>
<td>3.6 (3.2-3.9)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.8 (2.5-3.1)</td>
<td>2.8 (2.5-3.1)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>59</td>
<td>59</td>
<td>0.33</td>
</tr>
<tr>
<td>Female</td>
<td>40</td>
<td>40</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>3.5 (3.3-3.8)</td>
<td>2.7 (2.5-3.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.7 (3.3-4.2)</td>
<td>2.9 (2.5-3.2)</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>50</td>
<td>24</td>
<td>0.39</td>
</tr>
<tr>
<td>Past smoker</td>
<td>27</td>
<td>36</td>
<td>0.52</td>
</tr>
<tr>
<td>Current smoker</td>
<td>22</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.7 (3.4-4.1)</td>
<td>2.7 (2.4-3.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.6 (3.2-4.1)</td>
<td>2.8 (2.5-3.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.3 (2.9-3.7)</td>
<td>2.8 (2.5-3.1)</td>
<td></td>
</tr>
</tbody>
</table>

* Combined OMA score was defined as:
\[0.00425 \times \text{APE1} + 0.5419 \times \text{OGG1} - 0.02541 \times \text{MPG}.\]

One participant did not have known test values. This participant and the matched control were excluded from the analysis.

~ Of the 100 lung cancer cases, 30 had squamous cell carcinoma (SQCC), 46 had adenocarcinoma, 14 BAC, 4 adenosquamous carcinoma, 4 adenoBAC; 1 small cell carcinoma, and 1 unknown histology.

† Analysis of covariance comparing cases with controls, with matched pair and smoking status as a covariate. For smoking status subgroups, the covariates were age (continuous) and gender.

‡ Analysis of covariance comparing histological type within cases, with smoking status, age (continuous) and gender as covariates.

§ Test for interaction between case-control status and the variable of interest.

^ Analysis of covariance comparing subsets defined by the variable of interest and stratified by cases and controls, with smoking status, age (continuous) or gender as covariates, as appropriate.
Table 4. Conditional logistic regression analysis of combined OMA score based on OGG1, MPG and APE1 values in lung cancer patients and control subjects*

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. of control subjects (%)</th>
<th>No. of case patients (%)</th>
<th>Adjusted† OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score (per 1 SD decrease, 1.18U)‡</td>
<td>96 (100.0)</td>
<td>96 (100.0)</td>
<td>3.2 (1.7-5.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Cross-validation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score (per 1 SD decrease, 1.16 U)</td>
<td>96 (100.0)</td>
<td>96 (100.0)</td>
<td>2.7 (1.6-4.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>P&lt;0.001</td>
</tr>
<tr>
<td>Score (by tertiles)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;3.98</td>
<td>31 (32.3)</td>
<td>8 ( 8.3)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td>3.15–3.98</td>
<td>32 (33.3)</td>
<td>23 (24.0)</td>
<td>3.3 (1.1-9.5) P=0.03</td>
</tr>
<tr>
<td>≤3.14</td>
<td>33 (34.4)</td>
<td>65 (67.7)</td>
<td>9.7 (3.1-29.8) P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Trend test P&lt;0.001</td>
</tr>
<tr>
<td>Cross-validation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest tertile</td>
<td>30 (31.2)</td>
<td>12 (12.5)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td>Middle tertile</td>
<td>33 (34.4)</td>
<td>19 (19.8)</td>
<td>1.5 (0.6-3.7) P=0.42</td>
</tr>
<tr>
<td>Lowest tertile</td>
<td>33 (34.4)</td>
<td>65 (67.7)</td>
<td>5.6 (2.1-15.1) P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Trend test P&lt;0.001</td>
</tr>
</tbody>
</table>

* OR = odds ratio; CI = confidence interval.
† Conditional logistic regression for matched sets adjusted for smoking status (smoker, ex-smoker, never smoker). Three case patients did not have known smoking status and one control subject did not have known test values. These participants along with their matched subjects were excluded from the analysis.
‡ Score was defined as 0.00425×APE1 + 0.5419×OGG1 - 0.02541×MPG and was fitted in the conditional logistic regression model as a continuous variable. The value 1.18U is one standard deviation in the control group. The odds ratio for smoking, obtained with this model, was: ex-smoker v never smoker: 2.6 (95% CI= 1.0-6.3); current smoker v never smoker: 3.0 (1.1-8.2). The value 1.16U is one standard deviation of the cross-validated score in the control group.
|| Tertiles of the control subjects’ values. The upper tertile was chosen as the referent.
Figure 1
Sevilya et al 2013

A

OGG1

MPG
(AAG)

APE1
(APEX)
AP endonuclease

APE1

Pol β
XRCC1
DNA ligase III

B

Enzymatic activity, units/µg protein

10^7

10^6

10^5

10^4

10^3

10^2

10^1

10^0

OGG1

MPG

APE1

5.8
167
1.1x10^5
Low integrated DNA repair score and lung cancer risk


Cancer Prev Res Published OnlineFirst December 19, 2013.

Updated version Access the most recent version of this article at: doi:10.1158/1940-6207.CAPR-13-0318

Supplementary Material Access the most recent supplemental material at: http://cancerpreventionresearch.aacrjournals.org/content/suppl/2013/12/19/1940-6207.CAPR-13-0318.DC1

Author Manuscript Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.