Telomere length, arsenic exposure and risk of basal cell carcinoma of skin

Nalini Srinivas[1], Sivaramakrishna Rachakonda[1], Thomas Hielscher[2], Silvia Calderazzo[2], Peter Rudnai[3], Eugene Gurzau[4], Kvetoslava Koppova[5], Tony Fletcher[6], Rajiv Kumar[1,7*]

[1]Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg, Germany

[2]Department of Biostatistics, German Cancer Research Center, Heidelberg, Germany

[3]National Public Health Institute, Budapest, Hungary

[4]Environmental Health Center, Cluj, Romania

[5]Slovak Medical University Bratislava, Banska Bystrica

[6]London School of Hygiene and Tropical Medicine, London, UK

[7]German Consortium for Translational Research (DKTK), German Cancer Research Center, Heidelberg, Germany

Running title: Telomere length and basal cell carcinoma risk

Key words: telomere length, arsenic exposure, basal cell carcinoma

***Correspondence to:** Rajiv Kumar, Division of Molecular Genetic Epidemiology, German Cancer Research Center, Im Neunheimer Feld 580, 69120 Heidelberg, Germany

Email: r.kumar@dkfz.de

Abstract

Telomere length *per se* a heritable trait has been reported to be associated with different diseases including cancers. In this study, 528 cases with basal cell carcinoma of skin (BCC) and 533 healthy controls, we observed that decreased telomere length was associated with statistically significantly increased risk (OR 5.92, 95% CI 3.92-9.01, *P*<0.0001). We also observed that the higher arsenic exposure (>1.32 µg/L) was statistically significantly associated with decreased telomere length (β: -0.026, 95% CI: -0.05 to -0.003, *P* 0.02). The interaction between arsenic exposure and telomere length was statistically significant (*P* 0.02). Within each quantile of arsenic exposure, the individuals with shorter telomeres were at an increased risk of BCC, with highest risk being in the highest exposed group (OR 16.13, 95% CI 6.71-40.00, *P*<0.0001); followed by those in medium exposure group (OR 5.05, 95% CI 2.29-10.20, *P* <0.0001), and low exposure group (OR 3.44, 95% CI 1.81-6.54, *P* 0.0002). The combined effect of highest arsenic exposure and shortest telomeres on the risk of BCC (OR 10.56, 95% CI 5.14 to 21.70) showed statistically significant departure from additivity (interaction constant ratio 6.56, *P* 0.03). Our results show that in presence of the arsenic exposure, it is the decreased telomere length that predisposes individuals to increased risk of BCC with effect being synergistic in individuals with highest arsenic exposure and shortest telomeres.

Introduction

Basal cell carcinoma (BCC), which arises from the transformation of keratinocyte within the epidermis, accounts for 80-90% of all primary skin cancers (1). BCC are slow-growing and locally invasive tumors that result in extensive morbidity through recurrence and tissue destruction (2). The risk of BCC is associated with environmental factors such as ultraviolet (UV) exposure, ionizing radiation and arsenic exposure that induces DNA damage (3-7). BCC occurs mainly sporadically but some rare genetic disorders like Gorlin's syndrome and xeroderma pigmentosum result in multiple tumors with an early onset (8,9).

Telomeres at the chromosomal ends in humans consist of TTAGGG repeat sequences of approximately 10-15 kb of double stranded DNA ending in a single strand of up to 150-200 nucleotides. Telomere sequences together with associated proteins afford protection from end fusions to maintain genomic integrity (10-12). Telomere repeats, in the absence of telomerase in most of the somatic cells, progressively shorten by ~200 nucleotides with each cell division due to the 'end-replication problem' (13-15). Short or dysfunctional telomeres are recognized as DNA double stand breaks, triggering cells to undergo apoptosis or cellular senescence (16). Telomere length is considered as a potential biomarker of aging and age-related diseases (17). Studies conducted in large cohorts have reported that longer telomeres are associated with increased risk of different cancers, including melanoma, lung cancer, gliomas, lymphoma, and bladder neoplasms (18-21).

Telomere length is also influenced by a wide spectrum of genetic and environmental factors, which include different telomeropathies, electromagnetic radiation and arsenic exposure (22-25). Arsenic exposure induces oxidative stress leading to DNA damage, genomic instability, and shortened telomeres (26-28). Epidemiological studies indicate that long-term arsenic exposure in turn causes increased risk of cancers of the skin, bladder, lung and kidney (29-31). However, only a few studies are available to date on the effects of arsenic exposure on telomere length with limited subjects (26,32-34). In the present multi-center study, we found that BCC patients had statistically significantly shorter telomeres than controls. Arsenic exposure, a confounder in the study population, besides increasing the risk for BCC, was associated with statistically significant decreased telomere length. We observed a synergistic effect of arsenic exposure and short telomeres on the risk of BCC.

Methods

Study Population

This study constituted a part of a multinational European research project conducted in several regions of Hungary, Romania and Slovakia between 2002 and 2004. The cases and controls selected were of Hungarian, Romanian and Slovakian nationalities. On the basis of histopathological examinations by pathologists, 528 BCC cases and 533 hospital-based controls were included in the study, subject to fulfillment of a set of criteria. The blood samples were kept deep frozen at -80^{o}C until analysis. A general questionnaire that included cumulative sun exposure, effects of sun exposure on skin and age/s at diagnosis of basal cell carcinoma was completed by trained personnel after an interview of the recruited cases and controls. In addition, the interviews included details on demographic, life-style, socio-economic, medical history, occupational exposures, drinking and nutritional habits as well as detailed residential history (6,29,35,36). The concentration of arsenic in drinking water over the lifetime of an individual was measured in the study population. The assessment of arsenic exposure through drinking water has been described previously (37). Ethnic background of the cases and controls were recorded along with other characteristics of the study population. Local ethnic board approved the study plan and design.

Measurement of Relative Leukocyte Telomere Length

Relative leukocyte telomere length was measured in DNA from 533 healthy controls and 528 BCC cases using a multiplex quantitative real time PCR based assay, as described earlier by Cawthon et al., with minor modifications (38,39). Briefly, each reaction was performed in triplicates in optical 384-well reaction plates, in a 10µl volume, using 2µl of 5X HOT FIREPol probe qPCR Mix Plus with ROX (Solis BioDyne, Tartu, Estonia), 1.5µM of Syto-9 (Invitrogen, Thermo Fischer Scientific Inc., USA) and 5-10ng of genomic DNA. Non-template controls without any genomic DNA were included as negative controls. Four primers were used in each reaction to amplify telomere DNA (telg at 200nM and telc at 400nM) and the albumin gene (albugcr2 at 200nM and abldgcr2 at 400nM) and the primer sequences were designed accordingly **(Supplementary Table S1)**. Real-time PCR experiments were performed on a Viia-7 instrument (ABI, Applied Biosystems, Foster City, CA) using two simultaneous programs to acquire the respective C_T values for telomere sequence and albumin gene. The conditions for amplification of the telomere sequence were 95^{o}C/15 min, 2 cycles of

95°C/20s and 49°C/1 min followed by 27 cycles of 85°C/20s with signal acquisition at 59°C/30s. The conditions for albumin gene were 35 cycles of 95°C/15s, 85°C/30s with signal acquisition at 84°C/30s. The specificity for all amplifications was determined by melt-curve analysis performed at default settings (95°C/15s, 60°C/1min with continuous signal acquisition at 0.05°C/s ramping and 95°C/15s). Eight concentrations of a reference DNA sample (genomic DNA pooled from 17 healthy individuals) were included in triplicates in a 2-fold serial dilution (from 50ng to 0.4ng) to generate standard curves for telomere (T) and albumin (S) PCR products. The quality of the PCR amplification was performed using Applied Biosystems Viia-7 RUO software, version 1.2.2. The standard curve was used to quantify the telomere and albumin genes based on their respective C_T values and the obtained triplicate values were averaged. The relative telomere lengths were expressed as the ratio between T and S (T/S ratio). Inter-assay variation and intra-assay variation was determined by duplicating the reference DNA for all the dilutions in all the assays performed. The inter-assay coefficient of variation (CV) of the telomere and albumin assays were 1.52% and 0.90%, respectively and the intra-assay CV of the telomere and albumin assays were 0.63% and 0.67%, respectively. The PCR efficiencies of the telomere and albumin gene assays ranged between 98% and 103%.

Statistical analysis

Box plots were drawn to show the distribution of leukocyte telomere length (represented by T/S ratio) in BCC cases and controls and the differences were analyzed by two-tailed t-test. Scatter plots were drawn using telomere length as dependent variable and age as independent variable. Regression lines were drawn to summarize the correlation between the variables. The regression model that included the covariates was given by the equation: telomere length = ß0+ß1*age +ß2*BCC status +ß3*age*BCC status for each subject. The confidence intervals for beta coefficient/slope and mean differences between telomere length in cases and controls, and corresponding *P* values were estimated using generalized linear regression model by adjusting with sex as a co-variate.

The telomere length as continuous variable was log-transformed or used as categorical variable. As a categorical variable, the telomere length was grouped into three based on tertile distribution (q1: ≤ 0.49, q2: 0.50-0.73, q3: >0.73). The association of telomere length and other co-variates that included age, sex, eye color, country, complexion, sun effect,

arsenic exposure, rs861539 (*XRCC3*) polymorphism and *MC1R* polymorphism on BCC risk was estimated using univariate unconditional logistic regression. The risk estimates were also adjusted using a multivariate logistic regression and the corresponding odds ratios (OR) and 95% confidence intervals were determined.

Mendelian randomization (MR) was implemented to determine the causal effect of telomere length (exposure) on the risk of BCC (outcome) using genetic variants as instrumental variables. The Mendelian randomization was based on the following assumption: i) the genetic variants are associated with the exposure ii) the variants are independent of any confounders of the exposure-outcome association iii) the variants are not associated with the outcome conditional on the exposure and confounders of the exposure-outcome relationship (40). Mendelian randomization was carried out in R v3.4.3 (R project for statistical computing) using the R-package 'Mendelian Randomization' (41). In this model, the risk estimates were determined using Inverse-variance weighted (IVW) and maximum likelihood-based methods. Nine single nucleotide polymorphisms (SNPs) as valid instrumental variables were chosen from previously published GWAS (42-45). In a two-sample MR study, the ß-estimates and the standard errors of the instrument-exposure association were taken from GWAS (46). Association between the instrument-outcome was estimated using logistic regression at the allele levels in the study population. A one-sample MR study was conducted using non-overlapping datasets in the same population. Two datasets each were generated for cases and controls based on similar mean and variance for arsenic exposure. The association of the instrument-exposure was estimated using linear regression analysis in the control population of the first N/2 individuals (dataset 1). Linear regression model included telomere length as dependent variable with polymorphism, age (continuous), arsenic exposure (continuous), country and polymorphism*age as independent variables. The standardized beta estimates and standard error values for the variant allele of each polymorphism*age on telomere length was estimated. The association of the instrument-outcome was estimated using logistic regression, adjusting for the covariates-age, arsenic exposure and country, in the second N/2 individuals (dataset 2).

The values of arsenic exposure were not normally distributed and hence were log-transformed. Linear regression model was carried out to determine the effect modification of arsenic exposure on telomere length. Association between arsenic exposure and telomere

length was analyzed using telomere length as a dependent variable and log-transformed arsenic exposure, age and country as independent variables in the combined population. Arsenic exposure was taken as a continuous variable in the entire population and in individual categories based on the median distribution, with ≤1.32 µg/L as lower exposure group and >1.32 µg/L as higher exposure group. Arsenic exposure was also modeled as a continuous variable in categories according to the WHO's guideline for arsenic in drinking water (≤10 µg/L and >10 µg/L).

Further, the interaction between telomere length (log-transformed continuous variable) and arsenic exposure (log-transformed continuous variable) on BCC risk (as dependent variable) was assessed. Arsenic exposure was categorized based on tertile distribution (q1: ≤0.70, q2: 0.71-16.38, q3: >16.38). Univariate logistic regression analysis for telomere length (as continuous variable) on the risk of BCC was conducted in each of the three quantiles of arsenic exposure. Telomere length was adjusted for all the significant confounders in a multivariate regression analysis the on risk of BCC in each quantile of arsenic exposure.

Arsenic exposure-telomere length interaction on BCC risk was estimated by testing departure from additive or multiplicative risk-ratios. Based on tertile distribution, arsenic exposure was categorized as high exposure (HE), medium exposure (ME) and low exposure (LE) groups; and telomere length was categorized as long (L), medium (M) and short (S). The risk estimates were determined using logistic regression analysis adjusting for country and individuals age. Lowest exposure and longest telomere length, $R_{HE}R_L$ was used as reference to calculate the odds ratios for the remaining 8 groups. The estimates were calculated as $R_{HE}R_S$ for high exposure and short telomere length, $R_{HE}R_M$ for high exposure and medium telomere length, $R_{HE}R_L$ for high exposure and long telomere length; $R_{ME}R_S$ for medium exposure and short telomere length, $R_{ME}R_M$ for medium exposure and medium telomere length, $R_{ME}R_L$ for medium exposure and long telomere length; $R_{LE}R_S$ for low exposure and shorttelomere length and $R_{LE}R_M$ for low exposure and medium telomere length. Interaction constant ratio (ICR) was calculated to test departure from additivity (< or > 0) for high exposure group and short telomere length as (ICR_{high}= $R_{HE}R_S$ - $R_{HE}R_L$- $R_{LE}R_S$+1) and high exposure group and medium telomere length as (ICR_{high}= $R_{HE}R_M$ - $R_{HE}R_L$- $R_{LE}R_M$+1). Multiplicative interaction index (MII) was calculated to test departure from multiplicativity (< or > 1) for high exposure group and short telomere length as (MII_{high}= $R_{HE}R_S$ / ($R_{LE}R_S$*$R_{HE}R_L$)

and high exposure group and medium telomere length as ($ICR_{high}= R_{HE}R_M$ / ($R_{LE}R_M$*$R_{HE}R_L$). Similarly, ICR for medium exposure group and short telomere length was calculated as ($ICR_{med}= R_{ME}R_S - R_{ME}R_L - R_{LE}R_S+1$) and medium exposure group and medium telomere length as ($ICR_{med}= R_{ME}R_M - R_{ME}R_L - R_{LE}R_M+1$). MII for medium exposure group and shortest telomere length as ($MII_{med}= R_{ME}R_S$ / ($R_{LE}R_S$*$R_{ME}R_L$) and medium exposure group and medium telomere length as ($ICR_{med}= R_{ME}R_M$ / ($R_{LE}R_M$*$R_{ME}R_L$). Confidence intervals and *P*-values for ICR and MII were determined using bootstrap method with 10,000 simulations. All statistical analysis was performed using SAS, v.9.4 (SAS Institute, Inc., Cary, NC).

Results

Telomere length and BCC risk

The median age of 528 BCC cases at diagnosis was 67 years (IQR: 58-73 years) and that of 533 controls was 61 years (IQR: 52-70 years) with 236 (44.70 %) men and 292 (55.30 %) women in cases and 274 (51.40 %) men and 259 (48.60 %) women in controls. Relative telomere length was successfully measured in 524 cases and 527 controls by multiplex quantitative real-time PCR. The data analysis showed that the BCC cases carried statistically significantly shorter telomeres than the controls (T-test, *P* 2.5 x 10^{-20}). The median relative telomere length measured in the cases was 0.56 (IQR 0.44-0.68) and in the controls was 0.67 (IQR 0.56-0.78). A statistically significant inverse correlation (Pearson's correlation r -0.30, *P*<0.0001) was observed between telomere length and age. The telomere attrition per year was similar both in the cases (slope -0.005, 95% CI -0.006 to -0.003, *P*<0.0001) and the controls (slope -0.005, 95% CI -0.006 to -0.003, *P*<0.0001) and the mean difference in telomere attrition per year between the two groups, adjusted for sex, was not statistically significant (estimate 0.0001 95% CI -0.002 to 0.002, *P* 0.91) **(Figure 1)**.

A logistic regression model was used to determine the association between telomere length, as a continuous and categorical variable, and risk for BCC. The telomere data was transformed by taking the natural log of telomere length. On this scale, the effect of each unit decrease in ln(telomere length) was estimated to have an OR of 5.92 (95% CI 3.92 to 9.01, *P*<0.0001) on BCC risk. Analysis with telomere length as categorical variable based on tertile distribution showed that individuals in the first tertile (shortest telomeres) compared to those in the third tertile (longest telomeres) carried increased risk for BCC (OR 4.74, 95%CI 3.46 to 6.50), followed by individuals in second tertile (OR 2.06, 95%CI 1.52 to 2.80),

P<0.0001 **(Supplementary Table S2)**. Other factors associated with BCC risk included age, sex, eye colour, complexion, sun effect, country and log-transformed life-time arsenic exposure, carriers of rs861539 (*XRCC3*) polymorphism and *MC1R* polymorphisms **(Supplementary information and Supplementary Table S2)**. The effect of telomere length remained statistically significant after adjustment with all confounders (OR 4.57, 95% CI 2.82 to 7.41, *P*<0.0001) **(Table 1)**.

Mendelian Randomization

Mendelian randomization analysis was carried out to determine the causal effect of telomere length (exposure) on BCC risk (outcome) using single nucleotide polymorphisms as instrumental variables. The SNPs, represented by rs6060627, rs6772228, rs9257445, rs1317082, rs2487999, rs7726159, rs755017, rs412658, rs3027234, were associated with telomere length taken from previously published GWAS (42-45).

In a two-sample MR study, the association estimates (β-estimate in terms of kb of telomere length) and the standard error was determined for the variant allele of each SNP associated with 'short telomere length' from published GWAS. The estimates (ln(OR)) for the instrument-outcome association was calculated from logistic regression in this study **(Supplementary information, Supplementary Table S3, Supplementary Table S4)**. Analysis showed no statistically significant association between telomere length and risk of BCC using inverse variance weighted (IVW) method (OR 0.57, 95%CI 0.10 to 3.23, *P* 0.53) and maximum-likelihood method (OR 0.57, 95%CI 0.10 to 3.25, *P* 0.52). In addition, the *P* value for heterogeneity test statistic was statistically significant implying failure of instrumental variable assumptions for at least one of the SNPs (*P* 0.01 with 8 degrees of freedom) **(Supplementary Figure S1 and Supplementary Table S5)**. Since arsenic exposure was found to be a statistically significant confounder, we carried out Mendelian randomization in a one-sample study using non-overlapping datasets to determine the effect of telomere length on BCC risk (47).

For a non-overlapping sample setting, the cases and controls separately were split as two groups each (cases- dataset 1 and 2, controls- dataset 1 and 2) and having similar mean and variance for arsenic exposure. The cases have 261 individuals in both dataset 1 and dataset 2 and the controls have 265 individuals in dataset 1 and 264 in dataset 2. The median within the datasets of controls and cases were comparable **(Supplementary Table S6)**. The

estimates of instrument-outcome association were obtained from the logistic regression from dataset 2; the estimates of instrument-exposure association were obtained from the dataset 1 using control population **(Supplementary Table S7 and Supplementary Table S8)**. We independently assessed the validity of the instrumental variables assumptions using dataset 1 **(Supplementary information and Supplementary Table S9)**. The estimates were adjusted for the measured confounders in the study population that included age, arsenic exposure and country. The combined estimated magnitude of association of all the SNPs for a 1 standard deviation decrease in the telomere length was associated with an increased risk of BCC, OR of 1.83 (95% CI 0.98 to 3.40, *P* 0.06) using the IVW method; OR of 1.83 (95% CI 0.98 to 3.41, *P* 0.06) with the maximum-likelihood method, although not statistically significant. The *P* value for heterogeneity test statistic for the IVW and Maximum-likelihood method was not significant (*P* 0.10 with 9 degrees of freedom) **(Figure 2 and Supplementary Table S10)**.

Association between arsenic exposure and telomere length

Linear regression model in the combined population of cases and controls, adjusted for country and individuals age, showed a trend towards decreased telomere length with increase in the arsenic exposure (β: -0.006, 95% CI: -0.02 to -0.01, *P* 0.45). Stratification based on median distribution showed that the association of arsenic exposure on telomere length was statistically significant in the subgroup of individuals who were exposed to higher than median (>1.32 µg/L) measure of arsenic exposure with a β-estimate -0.026 (95% CI: -0.05 to -0.003, *P* 0.02) compared to those with lower than median exposure (≤ 1.32 µg/L; β: 0.026, 95% CI: -0.04 to 0.09, *P* 0.41). In a similar analysis using WHO's guideline for arsenic in drinking water (≤ 10µg/L), the estimates were found to be β: -0.043, 95% CI: -0.09 to -0.005, *P* 0.08 in the higher exposure group (>10 µg/L) and β: 0.011, 95% CI: -0.02 to 0.04, *P* 0.45 in the lower exposure group (≤ 10µg/L) **(Table 2)**.

Combined effect of arsenic exposure and telomere length on BCC risk

As statistically significant interaction was observed between arsenic exposure and telomere length on the risk of BCC in the study population (*P* 0.02), therefore, the data was analyzed for the effect of each unit decrease in ln(telomere length) on the risk of BCC within each of the three quantiles of arsenic exposure in the combined study population. The individuals with short telomeres in the highest exposure group of arsenic carried statistically

significantly increased risk of BCC, q3 (OR 16.13, 95% CI 6.71 to 40.00, *P*<0.0001), followed by those in medium exposure group, q2 (OR 5.05, 95% CI 2.29 to 10.20, *P* <0.0001), and in the low exposure group, q1 (OR 3.44, 95% CI 1.81 to 6.54, *P* 0.0002). The effect of telomere length on BCC risk within each quantile of arsenic exposure remained statistically significant after adjustment with confounders **(Table 3)**. The effect of arsenic exposure, adjusted for the unmeasured confounders between countries, on increased the risk of BCC has been previously reported on the present study population (29) **(Supplementary Table S2)**.

We also investigated the combined effect of arsenic exposure and telomere length on BCC risk, adjusted for age and country. The study population was stratified into nine sub-groups based on arsenic exposure and telomere length; three groups of arsenic exposure categorized as high, medium and low exposure and three groups of telomere length categorized as short, medium and long. The group with lowest arsenic exposure and longest telomere length was used as reference and the OR was determined for the remaining eight groups. The highest risk was observed for the group with shortest telomeres and highest exposure with an OR of 10.56 (95% CI 5.14 to 21.70). The interaction between highest arsenic exposure and shortest telomere length was statistically significantly greater than additive (ICR_{high} = 6.56, 95%CI= 0.72 to 17.90, P<0.03) and multiplicative interaction greater than 1, but not statistically significant (MII_{high} = 1.99, 95%CI= 0.65 to 5.98, P<0.23). The OR for the risk of BCC in the group with short telomeres and medium exposure was 3.12 (95% CI 1.73 to 5.62). Similarly, the group with medium telomeres and high exposure (OR 2.29, 95% CI 1.21 to 4.36) and the group with medium telomeres and medium exposure (OR 2.25, 95% CI 1.27 to 4.00) statistically significantly increased the risk of BCC. However, the additive and multiplicative interactions were not statistically significant in any of the remaining groups **(Figure 3 and Table 4)**.

Discussion

In the present multi-center based case-control study, we showed that individuals with short telomeres compared to the long telomeres were at a 5-fold increased risk for BCC. Increased arsenic exposure, besides an independent risk factor for BCC, showed a statistically significant association with decreased telomere length. The data analysis shows that arsenic in the group with arsenic exposure higher than the median (1.32 µg/L) was statistically

significantly associated with decreased telomere length. The effect of telomere length within each tertile of arsenic exposure showed that individuals with short telomeres in the highest exposure groups were at highest risk for BCC followed by those in low and medium exposure groups. We observed an 11-fold increased risk of BCC due to the combined effect of highest arsenic exposure and shortest telomere length when compared to the group with lowest exposure and longest telomeres. The notion was further augmented by the observation of greater than additive interaction between highest arsenic exposure and shortest telomere length in the study population.

Various polymorphisms shown to be associated with risk of different cancers have also shown to be associated with telomere length, with alleles associated with longer telomeres being associated with increased cancer risk (21,48,49). These findings are consistent with observational studies showing significant associations between increased telomere length and increased cancer risk (39,50,51). It was postulated that long telomeres increase the proliferative potential by delaying senescence (52). A Mendelian randomization study based on 3361 BCC cases and 11518 controls from the BCC case-control study nested within Nurse's Health study (NHS) and Health Professionals Follow-up study (HPFS), report that genetically increased telomere length increases the risk for BCC (21,53,54). Contrary to those reported observations, our data analysis showed an association between increased risk of BCC and short telomeres, due to arsenic exposure.

Mendelian randomization using a two-sample study failed to show the effect of short telomeres on BCC risk, which together with the instrument failure because of statistically significant heterogeneity probably indicated the observed effect in univariate and multivariate analysis was due to confounders, arsenic exposure in this case. That was further corroborated in one sample study that incorporated arsenic exposure into the model and showed increased risk of BCC due to decreased telomere length (55). A statistically significant heterogeneity in the two-sample study indicated that the samples used to estimate the SNP-outcome and SNP-exposure associations are not homogeneous, due to confounding in the exposure-outcome relationship across samples (56). Although one-sample MR study indicated an association between telomere length and risk of BCC, it has limitations from weak instrument bias and underestimation of true causal effect (56,57).

The direction of the effect of telomere length on BCC risk was mainly because of exposure to high concentrations of arsenic in our study population through drinking water. Our data showed that high arsenic exposure was statistically significantly associated with decreased telomere length. The relationship between arsenic exposure and telomere length has been assessed in both experimental and observational studies. One of the principle mechanisms of arsenic toxicity is oxidative stress (58). Experimental studies have shown that arsenic-induced oxidative stress leads to telomere attrition, DNA damage, chromosomal instability and apoptosis (22,59-61). *In vitro* studies of arsenic exposure suggests that lower concentrations of arsenic exposure (<1μM) either maintain or increases the telomere length, whereas, at higher concentrations (>1μM) there was a drastic decrease in the telomere length, supporting the observations from our study (62,63). Therefore, the direction of the effects of the exposure on telomere length is dependent on various environmental and genetic factors such as arsenic dose, duration of exposure, age and DNA repair mechanism (26).

Studies have indicated that arsenic tends to accumulate in the skin and is associated with BCC and other non-melanoma skin cancers (64-66). A cohort based study of Bangladeshi adult men and women reported an association between shorter telomere length and risk of arsenic-induced skin lesions, a similar observation as our present study (67). Our data showed that the combined effect of decreased telomere length and high arsenic exposure resulted in more than 10-fold increased risk of BCC with statistically significant departure from additivity. Previously, epidemiological studies have shown apparent synergistic interactions between arsenic and other risk factors of skin lesions such as smoking, fertilizer use and sunlight exposure (68,69). Our study is the first to show a synergistic effect of high arsenic exposure and short telomeres on the risk of BCC suggesting that arsenic exposure interacts with telomere length to further increase the risk of BCC.

In contrast to general observations of association between increased risk of cancers including BCC and increased telomere length as consistently reported in epidemiological studies, our data showed that the presence of arsenic exposure modulates the direction of the effect (21). We observed that in presence of arsenic exposure, it is the decreased telomere length that increased the risk of BCC with effect being synergistic in individuals with highest arsenic exposure and shortest telomeres.

References

1. Iwasaki, J.K., *et al.* (2012) The molecular genetics underlying basal cell carcinoma pathogenesis and links to targeted therapeutics. *J Am Acad Dermatol*, **66**, e167-78.
2. Stanoszek, L.M., *et al.* (2017) Histologic Mimics of Basal Cell Carcinoma. *Arch Pathol Lab Med*, **141**, 1490-1502.
3. Nan, H., *et al.* (2011) Genetic variants in telomere-maintaining genes and skin cancer risk. *Hum Genet*, **129**, 247-53.
4. Rochette, P.J., *et al.* (2010) Human telomeres are hypersensitive to UV-induced DNA Damage and refractory to repair. *PLoS Genet*, **6**, e1000926.
5. Pellegrini, C., *et al.* (2017) Understanding the Molecular Genetics of Basal Cell Carcinoma. *Int J Mol Sci*, **18**.
6. Surdu, S., *et al.* (2013) Occupational exposure to arsenic and risk of nonmelanoma skin cancer in a multinational European study. *Int J Cancer*, **133**, 2182-91.
7. Li, C., *et al.* (2016) Ionizing Radiation Exposure and Basal Cell Carcinoma Pathogenesis. *Radiat Res*, **185**, 217-28.
8. Didona, D., *et al.* (2018) Non Melanoma Skin Cancer Pathogenesis Overview. *Biomedicines*, **6**.
9. Rimkus, T.K., *et al.* (2016) Targeting the Sonic Hedgehog Signaling Pathway: Review of Smoothened and GLI Inhibitors. *Cancers (Basel)*, **8**.
10. Stewart, J.A., *et al.* (2012) Maintaining the end: roles of telomere proteins in end-protection, telomere replication and length regulation. *Mutat Res*, **730**, 12-9.
11. O'Sullivan, R.J., *et al.* (2010) Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol*, **11**, 171-81.
12. Heidenreich, B., *et al.* (2017) TERT promoter mutations in telomere biology. *Mutat Res*, **771**, 15-31.
13. d'Adda di Fagagna, F., *et al.* (2003) A DNA damage checkpoint response in telomere-initiated senescence. *Nature*, **426**, 194-8.
14. De Vitis, M., *et al.* (2018) Telomere Length Maintenance in Cancer: At the Crossroad between Telomerase and Alternative Lengthening of Telomeres (ALT). *Int J Mol Sci*, **19**.
15. Han, J., *et al.* (2009) A prospective study of telomere length and the risk of skin cancer. *J Invest Dermatol*, **129**, 415-21.
16. Kong, C.M., *et al.* (2013) Telomere shortening in human diseases. *FEBS J*, **280**, 3180-93.
17. Blackburn, E.H., *et al.* (2015) Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science*, **350**, 1193-8.
18. Walsh, K.M., *et al.* (2015) Longer genotypically-estimated leukocyte telomere length is associated with increased adult glioma risk. *Oncotarget*, **6**, 42468-77.
19. Iles, M.M., *et al.* (2014) The effect on melanoma risk of genes previously associated with telomere length. *J Natl Cancer Inst*, **106**.
20. Machiela, M.J., *et al.* (2016) Genetically predicted longer telomere length is associated with increased risk of B-cell lymphoma subtypes. *Hum Mol Genet*, **25**, 1663-76.
21. Haycock, P.C., *et al.* (2017) Association Between Telomere Length and Risk of Cancer and Non-Neoplastic Diseases: A Mendelian Randomization Study. *JAMA Oncol*, **3**, 636-651.
22. Mannan, T., *et al.* (2018) Associations of arsenic exposure with telomere length and naive T cells in childhood- A birth cohort study. *Toxicol Sci*.
23. Lustig, A., *et al.* (2016) Long term effects of radiation exposure on telomere lengths of leukocytes and its associated biomarkers among atomic-bomb survivors. *Oncotarget*, **7**, 38988-38998.
24. Ibanez-Cabellos, J.S., *et al.* (2018) Acute telomerase components depletion triggers oxidative stress as an early event previous to telomeric shortening. *Redox Biol*, **14**, 398-408.

25. Sarek, G., *et al.* (2015) Molecular basis of telomere dysfunction in human genetic diseases. *Nat Struct Mol Biol*, **22**, 867-74.
26. Borghini, A., *et al.* (2016) Arsenic exposure, genetic susceptibility and leukocyte telomere length in an Italian young adult population. *Mutagenesis*, **31**, 539-46.
27. Trusina, A. (2014) Stress induced telomere shortening: longer life with less mutations? *BMC Syst Biol*, **8**, 27.
28. Bhattacharjee, P., *et al.* (2013) Role of genomic instability in arsenic-induced carcinogenicity. A review. *Environ Int*, **53**, 29-40.
29. Leonardi, G., *et al.* (2012) Inorganic arsenic and basal cell carcinoma in areas of Hungary, Romania, and Slovakia: a case-control study. *Environ Health Perspect*, **120**, 721-6.
30. Garcia-Esquinas, E., *et al.* (2013) Arsenic exposure and cancer mortality in a US-based prospective cohort: the strong heart study. *Cancer Epidemiol Biomarkers Prev*, **22**, 1944-53.
31. Steinmaus, C., *et al.* (2014) Increased lung and bladder cancer incidence in adults after in utero and early-life arsenic exposure. *Cancer Epidemiol Biomarkers Prev*, **23**, 1529-38.
32. Li, H., *et al.* (2012) Arsenic exposure through drinking water is associated with longer telomeres in peripheral blood. *Chem Res Toxicol*, **25**, 2333-9.
33. Mo, J., *et al.* (2009) Elevated human telomerase reverse transcriptase gene expression in blood cells associated with chronic arsenic exposure in Inner Mongolia, China. *Environ Health Perspect*, **117**, 354-60.
34. Chatterjee, D., *et al.* (2015) Arsenic exposure through drinking water leads to senescence and alteration of telomere length in humans: A case-control study in West Bengal, India. *Mol Carcinog*, **54**, 800-9.
35. Lindberg, A.L., *et al.* (2006) Arsenic exposure in Hungary, Romania and Slovakia. *J Environ Monit*, **8**, 203-8.
36. Thirumaran, R.K., *et al.* (2006) Single nucleotide polymorphisms in DNA repair genes and basal cell carcinoma of skin. *Carcinogenesis*, **27**, 1676-81.
37. Engstrom, K.S., *et al.* (2015) Genetic variation in arsenic (+3 oxidation state) methyltransferase (AS3MT), arsenic metabolism and risk of basal cell carcinoma in a European population. *Environ Mol Mutagen*, **56**, 60-9.
38. Cawthon, R.M. (2009) Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res*, **37**, e21.
39. Rachakonda, S., *et al.* (2018) Telomere length, telomerase reverse transcriptase promoter mutations, and melanoma risk. *Genes Chromosomes Cancer*, **57**, 564-572.
40. Burgess, S., *et al.* (2017) Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology*, **28**, 30-42.
41. Yavorska, O.O., *et al.* (2017) MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *International Journal of Epidemiology*, **46**, 1734-1739.
42. Pooley, K.A., *et al.* (2013) A genome-wide association scan (GWAS) for mean telomere length within the COGS project: identified loci show little association with hormone-related cancer risk. *Hum Mol Genet*, **22**, 5056-64.
43. Codd, V., *et al.* (2013) Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet*, **45**, 422-7, 427e1-2.
44. Codd, V., *et al.* (2010) Common variants near TERC are associated with mean telomere length. *Nat Genet*, **42**, 197-9.
45. Mangino, M., *et al.* (2012) Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. *Hum Mol Genet*, **21**, 5385-94.
46. Aziz, N.A., *et al.* (2018) Telomere length as a modifier of age-at-onset in Huntington disease: a two-sample Mendelian randomization study. *J Neurol*.
47. Hartwig, F.P., *et al.* (2016) Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int J Epidemiol*, **45**, 1717-1726.

48. Zhang, C*., et al.* (2015) Genetic determinants of telomere length and risk of common cancers: a Mendelian randomization study. *Hum Mol Genet*, **24**, 5356-66.
49. Rode, L*., et al.* (2016) Long telomeres and cancer risk among 95 568 individuals from the general population. *Int J Epidemiol*, **45**, 1634-1643.
50. Zhang, X*., et al.* (2017) The Association of Telomere Length in Peripheral Blood Cells with Cancer Risk: A Systematic Review and Meta-analysis of Prospective Studies. *Cancer Epidemiol Biomarkers Prev*, **26**, 1381-1390.
51. Rachakonda, S*., et al.* (2018) Telomere length and survival in primary cutaneous melanoma patients. *Sci Rep*, **8**, 10947.
52. Willeit, P*., et al.* (2010) Telomere length and risk of incident cancer and cancer mortality. *JAMA*, **304**, 69-75.
53. Nan, H*., et al.* (2011) Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet*, **20**, 3718-24.
54. Zhang, M*., et al.* (2013) Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Hum Mol Genet*, **22**, 2948-59.
55. Sekula, P*., et al.* (2016) Mendelian Randomization as an Approach to Assess Causality Using Observational Data. *J Am Soc Nephrol*, **27**, 3253-3265.
56. Hemani, G*., et al.* (2018) Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet*, **27**, R195-R208.
57. Lawlor, D.A. (2016) Commentary: Two-sample Mendelian randomization: opportunities and challenges. *Int J Epidemiol*, **45**, 908-15.
58. Flora, S.J. (2011) Arsenic-induced oxidative stress and its reversibility. *Free Radic Biol Med*, **51**, 257-81.
59. Liu, L*., et al.* (2003) Oxidative stress contributes to arsenic-induced telomere attrition, chromosome instability, and apoptosis. *J Biol Chem*, **278**, 31998-2004.
60. Mo, J*., et al.* (2009) Elevated ERCC1 gene expression in blood cells associated with exposure to arsenic from drinking water in Inner Mongolia. *Anticancer Res*, **29**, 3253-9.
61. Mo, J*., et al.* (2006) Chronic arsenic exposure and oxidative stress: OGG1 expression and arsenic exposure, nail selenium, and skin hyperkeratosis in Inner Mongolia. *Environ Health Perspect*, **114**, 835-41.
62. Ferrario, D*., et al.* (2009) Arsenic induces telomerase expression and maintains telomere length in human cord blood cells. *Toxicology*, **260**, 132-41.
63. Zhang, T.C*., et al.* (2003) Effects of arsenic on telomerase and telomeres in relation to cell proliferation and apoptosis in human keratinocytes and leukemia cells in vitro. *Carcinogenesis*, **24**, 1811-7.
64. Mayer, J.E*., et al.* (2016) Arsenic and skin cancer in the USA: the current evidence regarding arsenic-contaminated drinking water. *Int J Dermatol*, **55**, e585-e591.
65. Martinez, V.D*., et al.* (2011) Arsenic exposure and the induction of human cancers. *J Toxicol*, **2011**, 431287.
66. Yu, H.S*., et al.* (2006) Arsenic carcinogenesis in the skin. *J Biomed Sci*, **13**, 657-66.
67. Zhang, C*., et al.* (2018) A study of telomere length, arsenic exposure, and arsenic toxicity in a Bangladeshi cohort. *Environ Res*, **164**, 346-355.
68. Chen, Y*., et al.* (2006) Modification of risk of arsenic-induced skin lesions by sunlight exposure, smoking, and occupational exposures in Bangladesh. *Epidemiology*, **17**, 459-67.
69. Melkonian, S*., et al.* (2011) A prospective study of the synergistic effects of arsenic exposure and smoking, sun exposure, fertilizer use, and pesticide use on risk of premalignant skin lesions in Bangladeshi men. *Am J Epidemiol*, **173**, 183-91.

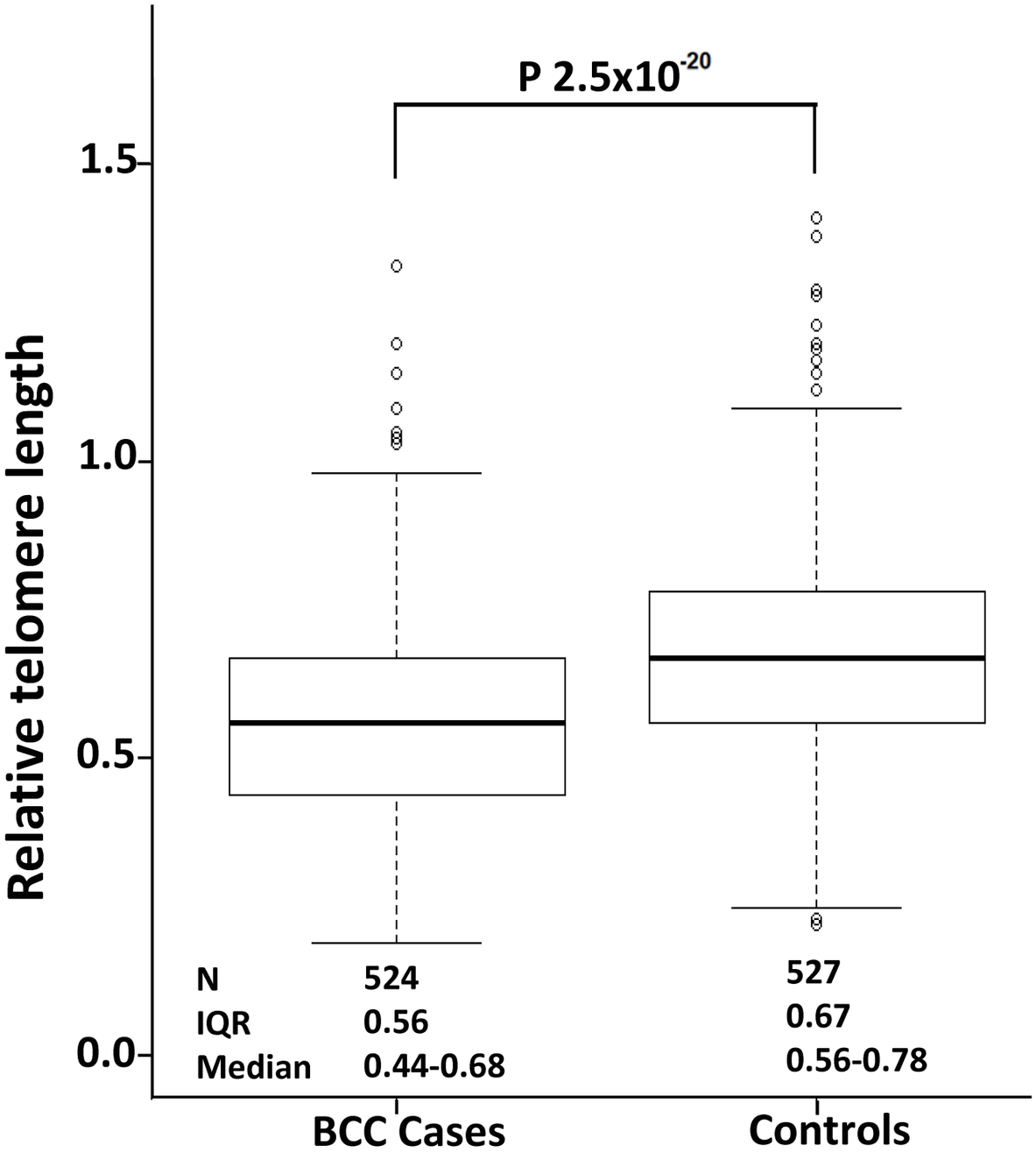
Legends to figures

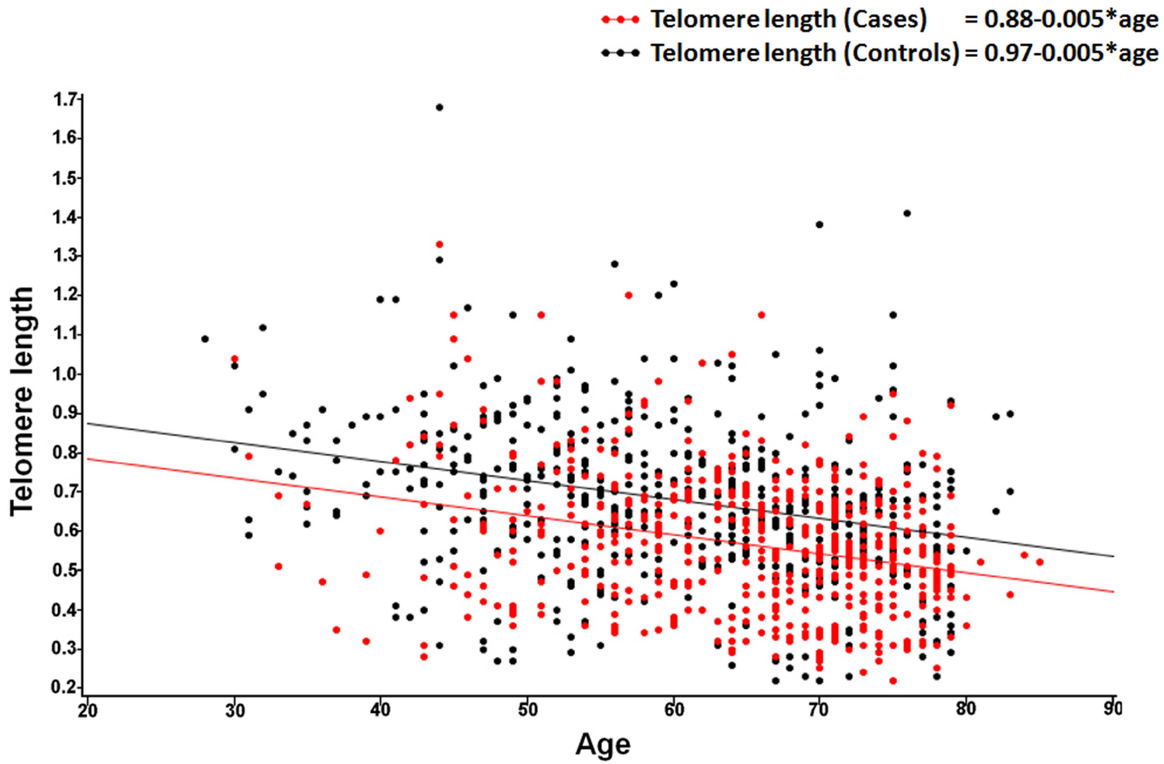
Figure 1: Distribution of relative telomere length in cases and controls. The box plots on the left depict differences in telomere lengths between BCC cases and controls. Number of cases and controls (N), median telomere length (median) and the corresponding inter-quantile range (IQR) are underneath the plots. Scatter plots on the right show relationship between telomere length and age in cases (red dots) and controls (black dots). Regression lines and corresponding equations are shown on the top right. Mean difference per year in telomere length between the cases and controls along with 95% confidence intervals are shown underneath the scatter plot.

Figure 2: One-sample Mendelian randomization using non-overlapping datasets showing association between telomere-length associated polymorphisms and risk of BCC. The scatter plots show the per-allele association with BCC risk plotted on the y-axes (represented as natural log of odds ratio) against the per-allele association with telomere length on x-axes from same study. The continuous blue line represents the Mendelian randomization estimate of telomere length on BCC risk. The model included all the nine SNPs chosen from GWAS and rs861539 (*XRCC3* polymorphism). The point estimate using the inverse-variance weighted method along with 95% confidence intervals are shown above the scatter plot.

Figure 3: Risk of BCC associated with interaction between arsenic exposure and telomere length. Arsenic exposure was categorized as high, medium and low exposure. Telomere length was categorized as short, medium and long. OR, adjusted for age and country, associated with the risk due to various combinations are given on top of bars and those that are significant are indicated by bold font. Lowest exposure and longest telomere group was used as reference.

Figure 1





Groups	Telomere length-mean difference/year[a]	95% CI	*P*-value
BCC Cases-Controls	0.0001	-0.002 to 0.002	0.91
[a]Mean telomere length was adjusted for age and sex			

Figure 2

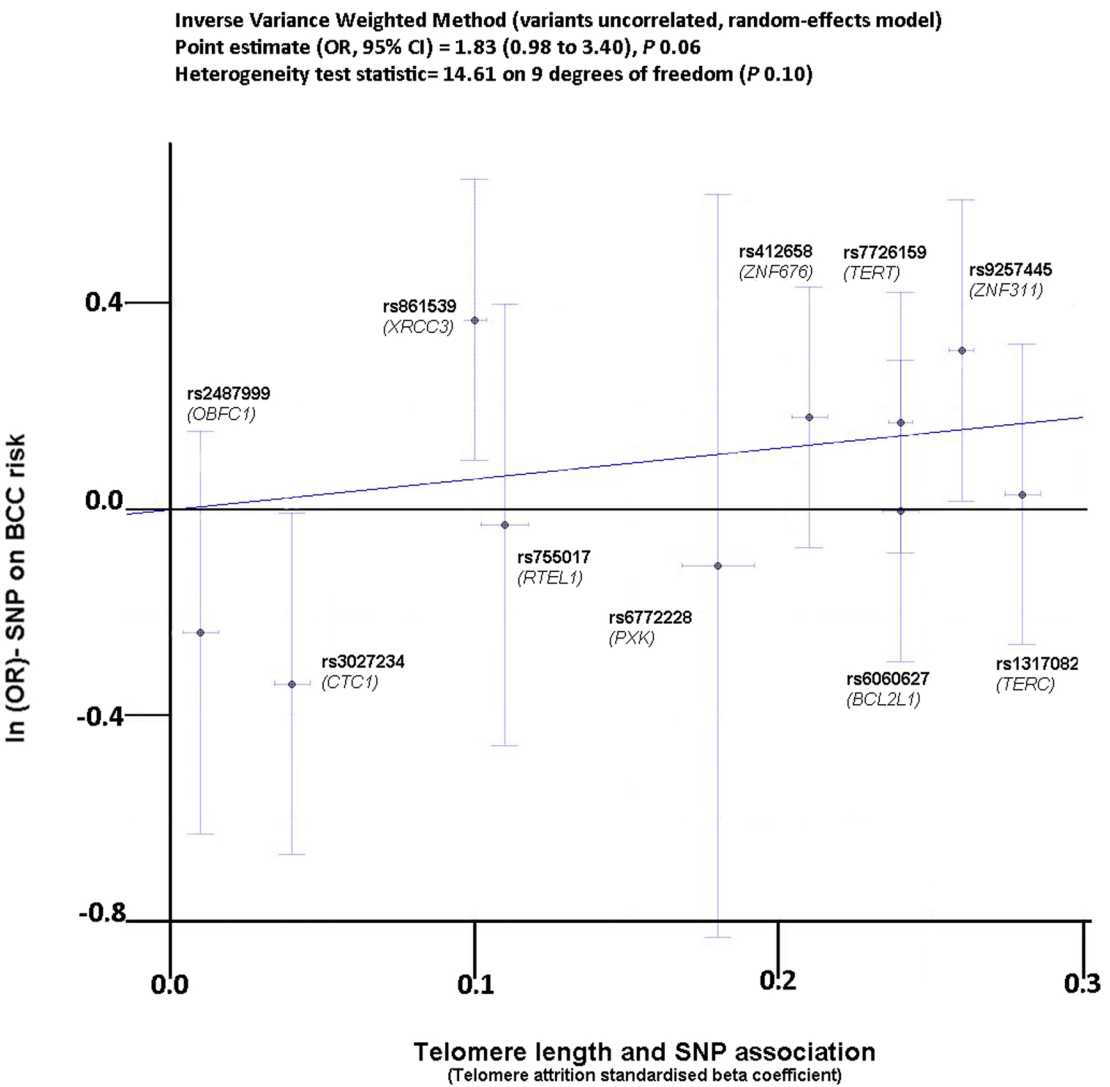

Inverse Variance Weighted Method (variants uncorrelated, random-effects model)
Point estimate (OR, 95% CI) = 1.83 (0.98 to 3.40), P 0.06
Heterogeneity test statistic= 14.61 on 9 degrees of freedom (P 0.10)
rs412658 (ZNF676)
rs7726159 (TERT)
rs9257445 (ZNF311)
rs861539 (XRCC3)
rs2487999 (OBFC1)
rs755017 (RTEL1)
rs6772228 (PXK)
rs3027234 (CTC1)
rs6060627 (BCL2L1)
rs1317082 (TERC)
0.4
0.0
-0.4
-0.8
0.0
0.1
0.2
0.3
ln (OR)- SNP on BCC risk
Telomere length and SNP association
(Telomere attrition standardised beta coefficient)

Figure 3

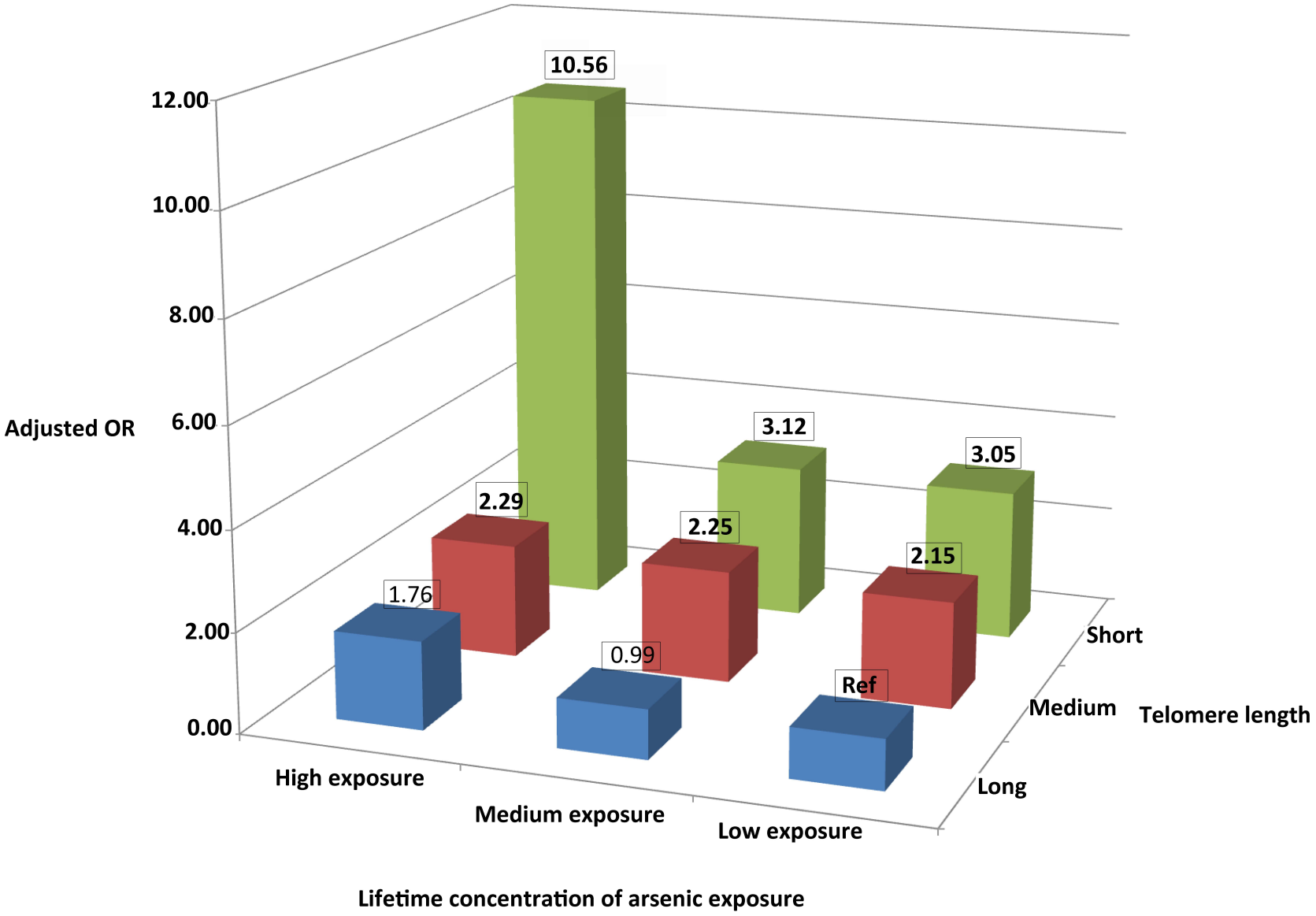

Adjusted OR
12.00
10.00
8.00
6.00
4.00
2.00
0.00
10.56
3.12
3.05
2.29
2.25
2.15
1.76
0.99
Ref
High exposure
Medium exposure
Low exposure
Lifetime concentration of arsenic exposure
Short
Medium
Long
Telomere length

Table 1. Multivariate Analysis on the effect of telomere length and study-specific covariates

	Type	Cases	Controls	OR (95% CI)	*P*-value[b]
Telomere length[a] (log transformed)	*Continuous*	480	490	**4.57 (2.82-7.41)**	**<0.0001**
Age	*Continuous*	480	490	**1.03 (1.02-1.04)**	**<0.0001**
Sex	Male Female	217 263	251 239	**Ref** 1.31 (0.99-1.74)	0.06
Eye Colour	Blue/Green Hazel/Brown Others	290 164 26	260 201 29	1.26 (0.93-1.70) **Ref** 0.81 (0.43-1.53)	0.18
Country	Hungary Romania Slovakia	145 129 206	230 129 131	**Ref** **2.50 (1.53-4.09)** **3.48 (2.20-5.51)**	**<0.0001**
Lifetime Concentration (Arsenic exposure in µg/L) (log transformed)	q1 (≤0.70) q2 (0.71-16.38) q3 (>16.38)	166 166 148	150 165 175	**Ref** 1.16 (0.81-1.67) **1.82 (1.11-2.99)**	0.06
Complexion	Light Medium/Dark	257 223	198 292	**1.69 (1.25-2.28)** **Ref**	**0.0007**
Sun effect	Blistered/burnt Mild burn Tan/No effect	171 161 148	135 148 207	**1.61 (1.13-2.29)** **1.60 (1.13-2.26)** **Ref**	**0.009**
rs861539 *(XRCC3)*	Non-carriers (CC genotype) Carriers (CT,TT genotype)	206 274	164 326	**Ref** **0.68 (0.51-0.90)**	**0.001**
MC1R[c]	Non-carriers Carriers	148 332	216 274	**Ref** **1.58 (1.18-2.11)**	**0.002**

[a]Odds ratio, OR calculated for each unit decrease in log-transformed telomere length.

[b]*P* -values were derived from χ^2-test, two sided and considered statistically significant if < 0.05, indicated by bold font.

[c]The individuals with any one *MC1R* polymorphism, that include V60L, D84E, V92M, R142H, R151C, I155T, R160W, R163Q, D294H, and T314T, are considered as carriers.

Table 2. Association between arsenic exposure and telomere length

Life time Concentration[a] **(Arsenic Exposure in µg/L)**	**Combined population**[b]		
	N	**β-estimate (95% CI)**	***P*-value**[c]
Continuous	1041	-0.006 (-0.02- 0.01)	0.45
Categorical:			
>1.32 µg/L	519	**-0.026 (-0.05- -0.003)**	**0.03**
≤1.32 µg/L	522	0.027 (-0.04- 0.09)	0.40
>10 µg/L	317	-0.043 (-0.09- 0.005)	0.08
≤10 µg/L	724	0.011 (-0.02- 0.04)	0.45

[a] Lifetime concentration was log-transformed and taken as a continuous variable and categorised as dichotomous variable based on median (1.32µg/L) and according to WHO guidelines for arsenic in drinking water (10µg/L).
[b] β-estimates adjusted for age and country.
[c] *P* -values were derived from χ^2-test, two sided and considered statistically significant if ≤ 0.05, indicated by bold font.

Table 3. Effect of telomere length as a continuous variable on BCC risk in each tertile of lifetime concentration of arsenic exposure

LTC exposure[a] (log-transformed)	Telomere length[b] (Median, IQR) (log-transformed)	Univariate Analysis				Multivariate Analysis[e]			
		Cases	Controls	OR (95% CI)[c]	P-value[d]	Cases	Controls	OR (95% CI)[c]	P-value[d]
q1 (≤ 0.70)	Median: 0.59 (IQR: 0.44-0.72)	183	163	**3.44 (1.81-6.54)**	**0.0002**	166	150	**3.24 (1.54-6.80)**	**0.002**
q2 (0.71-16.38)	Median: 0.63 (IQR: 0.46-0.71)	172	175	**5.05 (2.29-10.20)**	**<0.0001**	166	165	**2.62 (1.12-6.13)**	**0.03**
q3 (>16.38)	Median: 0.64 (IQR: 0.53-0.74)	163	185	**16.13 (6.71-40.00)**	**<0.0001**	148	175	**20.00 (6.71-58.82)**	**<0.0001**

[a] Lifetime concentration of arsenic exposure (LTC) was log transformed and grouped into tertiles.

[b] Telomere length was log-transformed and taken as a continuous variable for each group of LTC exposure.

[c] Odds ratio calculated for each unit decrease in log-transformed telomere length.

[d] *P* -values were derived from χ^2-test, two sided and considered statistically significant if ≤ 0.05, indicated by bold font.

[e] Odds ratio in multivariate model was adjusted for age, sex, eye colour, country, complexion, suneffect, *XRCC3* polymorphism, *MC1R* polymorphisms.

Table 4. Effect of interaction between arsenic exposure and telomere length on risk of BCC

Telomere length	Lifetime concentration of arsenic exposure measured in µg/L								
	Lowest exposure Median (IQR): 0.61 (0.48-0.70)			**Medium exposure** Median (IQR): 1.32 (0.93-2.75)			**Highest exposure** Median (IQR): 26.02 (16.38-41.90)		
	Cases	**Controls**	**OR (95%CI)**	**Cases**	**Controls**	**OR (95%CI)**	**Cases**	**Controls**	**OR (95%CI)**
Longest Median (IQR): 0.78 (0.73-0.89)	36	71	**Reference**	41	88	0.99 (0.56 to 1.75)	41	92	1.76 (0.93 to 3.35)
Medium Median (IQR): 0.61 (0.57-0.65)	57	45	**2.15 (1.21 to 3.82)**	60	51	**2.25 (1.27 to 4.00)**	50	72	**2.29 (1.21 to 4.36)**
Interaction with arsenic exposure [b]									
MII (95% CI)				1.06 (0.38 to 2.76)		*P* < 0.90	0.60 (0.21 to 1.76)		*P* < 0.37
ICR (95% CI)				0.08 (-2.04 to 2.11)		*P* < 0.93	-0.69 (-3.10 to 1.72)		*P* < 0.54
Shortest Median (IQR): 0.43 (0.35-0.49)	90	47	**3.05 (1.76 to 5.28)**	71	36	**3.12 (1.73 to 5.62)**	72	21	**10.56 (5.14 to 21.70)**
Interaction with arsenic exposure									
MII (95% CI)				1.04 (0.38 to 2.76)		*P* < 0.93	**1.99 (0.65 to 5.98)**		***P* < 0.23**
ICR (95% CI)				0.05 (-2.66 to 2.84)		*P* < 0.98	**6.56 (0.72 to 17.90)**		***P* < 0.03**

[a] OR and 95% CI for the risk associated with combined effect of arsenic exposure and telomere length on BCC risk was adjusted for age and country. Global p-value for the combined effect was <0.0001.

[b] MII is multiplicative interaction index and ICR, interaction constant ratio. The lowest arsenic exposure and longest telomeres was used as a single reference group for the interaction analysis.