OPEN ACCESS

Phenotypic and genetically predicted leucocyte telomere length and lung cancer risk in the prospective UK Biobank

Jason YY Wong ,[1] Batel Blechter,[2] Aubrey K Hubbard,[2] Mitchell J Machiela,[2] Jianxin Shi,[2] Shahinaz M Gadalla,[2] Wei Hu,[2] Mohammad L Rahman ,[2] Nathaniel Rothman,[2] Qing Lan[2]

[1]Epidemiology and Community Health Branch, National Heart Lung and Blood Institute, Bethesda, Maryland, USA
[2]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland, USA

Correspondence to
Dr Jason YY Wong, Epidemiology and Community Health Branch, National Heart Lung and Blood Institute, Bethesda, Maryland, USA; jason.wong@nih.gov
Published Online First 17 January 2024

ABSTRACT

We investigated phenotypic leucocyte telomere length (LTL), genetically predicted LTL (gTL), and lung cancer risk among 371 890 participants, including 2829 incident cases, from the UK Biobank. Using multivariable Cox regression, we found dose-response relationships between longer phenotypic LTL (p-trend$_{continuous}$=2.6×10^{-5}), longer gTL predicted using a polygenic score with 130 genetic instruments (p-trend$_{continuous}$=4.2×10^{-10}), and overall lung cancer risk, particularly for adenocarcinoma. The associations were prominent among never smokers. Mendelian Randomization analyses supported causal associations between longer telomere length and lung cancer (HR$_{per\ 1\ SD\ gTL}$=1.87, 95% CI: 1.49 to 2.36, p=4.0×10^{-7}), particularly adenocarcinoma (HR$_{per\ 1\ SD\ gTL}$=2.45, 95%CI: 1.69 to 3.57, p=6.5×10^{-6}).

INTRODUCTION

Telomere length reflects the cumulative burden of exposures, endogenous factors, and age. We investigated phenotypic leucocyte telomere length (LTL), telomere length predicted using a polygenic score (PGS; gTL), and lung cancer risk in the United Kingdom (UK) Biobank. Further, we examined the associations by histology and subgroups defined by smoking status and sex. To evaluate causal relationships, we conducted Mendelian Randomization (MR) analyses using 130 genetic instruments that predict longer LTL.[1 2] Our study could help identify high-risk subpopulations that do not have but may benefit from lung cancer screening.

MATERIALS AND METHODS

Study population

The UK Biobank is described previously[3] and in the *Supplementary Materials*. Among 502 409 participants at baseline, we excluded 372 subjects with discrepancies between self-reported and genetic sex; 46 577 with any cancer diagnosis; 57 253 with respiratory diseases; and 4779 with hematologic/immunologic disorders. Our analytic dataset had 393 428 participants.

Follow-up time started at the visit date to the assessment centres in 2006–2010 and ended at the date of primary incident lung cancer diagnosis, death, or administrative censoring. Lung cancer diagnosis was defined by International Classification of Diseases 10th revision (ICD-10) codes C34.0-C34.9. ICD-O-3 code 8140 defined adenocarcinoma (LUAD), while 8052, 8084, 8073, and 8083 defined squamous cell carcinoma (SCC).

LTL measurements

The DNA extraction, multiplex quantitative polymerase chain reaction (qPCR) assay, and quality control procedures were previously described.[4] The ratio of telomere (T) to single-copy gene (S) copy number (*'T/S ratio'*) reflects the average telomere abundance across all chromosomes in leucocytes of an individual, which was further adjusted for batch variation (*'adjusted relative* T/S ratio').[4]

PGS for longer gTL and MR analyses

Using a one-sample approach,[1] we constructed a weighted PGS that predicted longer gTL with 130 single nucleotide polymorphisms (SNPs) identified among Europeans in the UK Biobank. These SNPs explained 4.54% of the variance in LTL.[1] We note that LTL is affected by the exposome, which can influence LTL variation but unlikely bias the identification of LTL-related SNPs. The SNPs were linkage disequilibrium pruned ($r^2 > 0.01$), reached genome-wide significance ($p<8.31\times10^{-9}$), and included in previous MR analyses.[1] The weights were based on reported Z-standardised beta-estimates.[1] To estimate causal associations between gTL and lung cancer and its subtypes among Europeans, we conducted MR analyses using the same 130 SNPs with MR-PRESSO in R. The HR from MR analyses reflect increased lung cancer risk per one SD increase in gTL.[1]

Statistical analyses

Cox regression was used to estimate HRs and 95% CI of incident lung cancer in relation to quartiles (Q) of phenotypic LTL ('*adjusted relative* T/S Ratio'; Q1:<0.74; Q2:0.74 to<0.82; Q3:0.82 to<0.91; Q4:≥0.91), adjusted for age, sex, race/ethnicity, detailed smoking history/intensity, assessment centre, body mass index, Townsend Deprivation Index, alcohol intake, secondhand smoke exposure, and leucocyte differentials. Follow-up time was the timescale. LUAD and SCC were analysed separately. We fitted separate Cox models with gTL as the main effect among Europeans, without leucocyte differential adjustment.

Linear trends were estimated using continuous gTL and log-transformed phenotypic LTL.

To cite: Wong JYY, Blechter B, Hubbard AK, *et al*. *Thorax* 2024;**79**:274–278.

Table 1 Associations between measured phenotypic leucocyte telomere length and lung cancer risk in 371 890 participants from the UK Biobank

Quartile	Adjusted relative T/S Ratio Cutoffs	No. of incident cases	HR	95% CI lower	95% CI upper	p-value	
		Overall lung cancer (2829 cases)					
1	<0.74	812	1.00				
2	0.74 to<0.82	666	1.00	0.90	1.11	0.97	
3	0.82 to<0.91	626	1.08	0.97	1.19	0.18	
4	≥0.91	612	1.27	1.15	1.42	8.9×10^{-6}	*
					p-trend	2.6×10^{-5}	*
		II) Lung adenocarcinoma (1078 cases)					
		No. of incident cases	**HR**	**95% CI Lower**	**95% CI Upper**	**p-value**	
1	<0.74	254	1.00				
2	0.74 to<0.82	256	1.19	1.00	1.41	0.06	
3	0.82 to<0.91	237	1.23	1.03	1.47	0.02	*
4	≥0.91	292	1.78	1.50	2.12	4.0×10^{-11}	*
					p-trend	6.6×10^{-10}	*
		III) Lung squamous cell carcinoma (487 cases)					
		No. of incident cases	**HR**	**95% CI Lower**	**95% CI Upper**	**p-value**	
1	<0.74	164	1.00				
2	0.74 to<0.82	118	0.95	0.75	1.21	0.68	
3	0.82 to<0.91	111	1.06	0.83	1.35	0.66	
4	≥0.91	72	0.88	0.67	1.17	0.38	
					p-trend	0.48	

Multivariable Cox regression models were used to estimate HR and 95% CI of incident lung cancer in relation to quartiles of *adjusted relative* T/S Ratio (ie, measured phenotypic LTL), adjusted for age at recruitment (continuous), sex (men vs women), race/ethnicity/ancestry (reference: European), detailed smoking history/intensity (28 categories, reference: never smokers), study assessment centre, BMI (reference: ≥18.5 to <25.0 kg/m^2), Townsend Deprivation Index (continuous), alcohol intake (reference: never drinker), exposure to secondhand smoke (ever vs never), and leucocyte subtype differentials (% lymphocytes, neutrophils, eosinophils, basophils and monocytes). P-trends were estimated using continuous log-transformed *adjusted relative* T/S ratio.

*P-values<0.05 were considered statistically significant. A total of 371 890 participants with complete data on independent variables were included in the analyses. Discrepancy in counts was due to missing LTL data.

Stratified analyses were conducted among subgroups defined by smoking status (never, former, current) and sex. We tested for multiplicative effect modification of LTL-lung cancer associations by smoking status and sex using cross-product terms. Heterogeneity between HR estimates was assessed using Z-score sign tests.

RESULTS

Phenotypic LTL and lung cancer

There were 2829 incident lung cancer cases, including 1078 LUAD and 487 SCC, diagnosed over the 12.36±1.64 SD year follow-up. The average time between LTL measurements and lung cancer diagnosis was 6.54±3.22 SD years. We found a strong dose-response relationship between longer phenotypic LTL and increased lung cancer risk (p-trend$_{continuous}=2.6\times10^{-5}$), particularly LUAD (p-trend$_{continuous}=6.6\times10^{-10}$) but not SCC (table 1; p-difference$_{LUAD\ vs.\ SCC}=6.7\times10^{-5}$). Findings were similar among Europeans and when including participants with chronic respiratory diseases at enrollment (online supplemental table 1).

We observed multiplicative effect modification of the LTL-lung cancer association by smoking, with attenuation among former (p-interaction=0.008) and current smokers (p-interaction=0.008) vs never smokers. We did not detect an interaction between phenotypic LTL and sex (p-interaction=0.909). Among subgroups, we found associations between longer phenotypic LTL and increased lung cancer risk among never smokers (p-trend=3.3×10^{-7}), including never-smoking women (p-trend=4.0×10^{-5}) and men (p-trend=2.0×10^{-3}), as well as former-smoking women (p-trend=5.7×10^{-4}) (table 2). When analysing LUAD, the trends were consistent with overall lung cancer (online supplemental table 2). Due to sparse data, we could not analyse SCC among never smokers and consistent associations were not detected in other subgroups (online supplemental table 2)

PGS and MR analyses among Europeans

In PGS analyses, we found strong dose-response relationships between longer gTL and lung cancer risk (p-trend=4.2×10^{-10}), particularly LUAD (p-trend=2.1×10^{-8}) but not SCC (figure 1). We did not detect multiplicative interactions between gTL and smoking or sex (p-interactions>0.05).

In MR analyses, we found evidence for causal associations between longer gTL and lung cancer ($HR_{per\ 1\ SD\ gTL}=1.87$, 95% CI: 1.49 to 2.36, p=4.0×10^{-7}), particularly LUAD ($HR_{per\ 1\ SD\ gTL}=2.45$, 95% CI: 1.69 to 3.57, p=6.5×10^{-6}) but not SCC ($HR_{per\ 1\ SD\ gTL}=1.19$, 95% CI: 0.74 to 1.93, p=0.47; p-difference$_{LUAD\ vs.\ SCC}=0.02$). No horizontal pleiotropy was detected for any SNP.

DISCUSSION

We confirm and expand on previous studies with important findings among subpopulations defined by smoking and sex.

Table 2 Associations between measured phenotypic leucocyte telomere length and overall lung cancer risk in the UK Biobank among subgroups defined by sex and smoking status

Quartile	Adjusted relative T/S Ratio Cutoffs	No. of incident cases	HR	95% CI lower	95% CI upper	p-value	
I) NEVER SMOKERS		424 cases / 205 237 subjects					
1	<0.74	91	1.00				
2	0.74 to<0.82	74	0.85	0.63	1.16	0.31	
3	0.82 to<0.91	109	1.33	1.00	1.75	4.9×10^{-2}	*
4	≥0.91	130	1.72	1.31	2.26	9.9×10^{-5}	*
					p-trend	3.3×10^{-7}	*
II) FORMER SMOKERS		1254 cases / 125 995 subjects					
1	<0.74	352	1.00				
2	0.74 to<0.82	295	1.01	0.87	1.18	0.88	
3	0.82 to<0.91	283	1.13	0.96	1.32	0.14	
4	≥0.91	272	1.33	1.13	1.56	5.3×10^{-4}	*
					p-trend	3.0×10^{-3}	*
III) CURRENT SMOKERS							
		1120 cases / 38 808 subjects					
1	<0.74	360	1.00				
2	0.74 to<0.82	288	1.05	0.90	1.23	0.53	
3	0.82 to<0.91	227	0.96	0.81	1.13	0.61	
4	≥0.91	206	1.09	0.92	1.30	0.33	
					p-trend	0.49	
IV) NEVER SMOKING WOMEN		271 cases / 118 044 subjects					
Quartile	Adjusted Relative T/S Ratio Cutoffs	No. of incident cases	HR	95% CI Lower	95% CI Upper	p-value	
1	<0.74	58	1.00				
2	0.74 to<0.82	37	0.62	0.41	0.94	0.02	*
3	0.82 to<0.91	70	1.18	0.83	1.67	0.36	
4	≥0.91	91	1.55	1.11	2.17	0.01	*
					p-trend	4.0×10^{-5}	*
V) NEVER SMOKING MEN		153 cases / 87 193 subjects					
1	<0.74	33	1.00				
2	0.74 to<0.82	37	1.33	0.83	2.12	0.24	
3	0.82 to<0.91	39	1.60	1.00	2.55	0.05	
4	≥0.91	39	2.03	1.27	3.26	3.3×10^{-3}	*
					p-trend	2.0×10^{-3}	*
VI) FORMER SMOKING WOMEN		545 cases / 59 763 subjects					
1	<0.74	114	1.00				
2	0.74 to<0.82	120	1.09	0.85	1.41	0.50	
3	0.82 to<0.91	135	1.30	1.01	1.67	0.04	*
4	≥0.91	147	1.57	1.23	2.02	3.3×10^{-4}	*
					p-trend	5.7×10^{-4}	*
VII) FORMER SMOKING MEN		709 cases / 66 232 subjects					
1	<0.74	238	1.00				

Continued

Table 2 Continued

Quartile	Adjusted Relative T/S Ratio Cutoffs	No. of incident cases	HR	95% CI Lower	95% CI Upper	p-value	
2	0.74 to<0.82	175	0.99	0.81	1.20	0.89	
3	0.82 to<0.91	148	1.03	0.84	1.27	0.75	
4	≥0.91	125	1.19	0.96	1.48	0.12	
					p-trend	0.32	
VIII) CURRENT SMOKING WOMEN		455 cases / 16 649 subjects					
1	<0.74	145	1.00				
2	0.74 to<0.82	106	0.81	0.63	1.04	0.09	
3	0.82 to<0.91	93	0.73	0.56	0.95	0.02	*
4	≥0.91	99	0.92	0.71	1.20	0.55	
					p-trend	0.41	
IX) CURRENT SMOKING MEN		665 cases / 22 159 subjects					
1	<0.74	215	1.00				
2	0.74 to<0.82	182	1.23	1.01	1.50	0.04	*
3	0.82 to<0.91	134	1.13	0.91	1.41	0.26	
4	≥0.91	107	1.19	0.94	1.51	0.15	
					p-trend	0.14	

Multivariable Cox regression models were used to estimate HR and 95% CI of incident overall lung cancer in relation to quartiles of *adjusted relative* T/S Ratio (ie, measured phenotypic LTL), adjusted for age at recruitment (continuous), race/ethnicity/ancestry (reference: European), detailed smoking history/intensity when analysing former and current smokers, study assessment centre, BMI (reference: ≥18.5 to <25.0 kg/m^2), Townsend Deprivation Index (continuous), alcohol intake (reference: never drinker), exposure to secondhand smoke (ever vs never), and leucocyte subtype differentials (% lymphocytes, neutrophils, eosinophils, basophils and monocytes). P-trends were estimated using continuous log-transformed *adjusted relative* T/S ratio.
*P-values<0.05 were considered statistically significant. Discrepancy in counts were due to missing LTL data.

Previously, we reported associations between longer phenotypic LTL and increased lung cancer risk when pooling three nested case-control studies, particularly for LUAD among Europeans.[5] Those findings were corroborated in separate studies of Chinese[6] and heavy-smoking European populations.[7] Our PGS and MR analyses were in agreement with studies that found associations between longer gTL and increased lung cancer risk.[8–10] In particular, our MR analyses support causal associations between gTL and lung cancer. The magnitude of our gTL and phenotypic LTL effects were similar; suggesting that the genetic component of LTL captured by gTL is an effective surrogate marker, despite explaining a small proportion of variance. Our findings support longer pre-diagnostic LTL in processes relevant to lung carcinogenesis, prominently among non-smokers. Longer LTL potentially reflects decreased senescence or higher replicative potential in pre-cancerous cells,

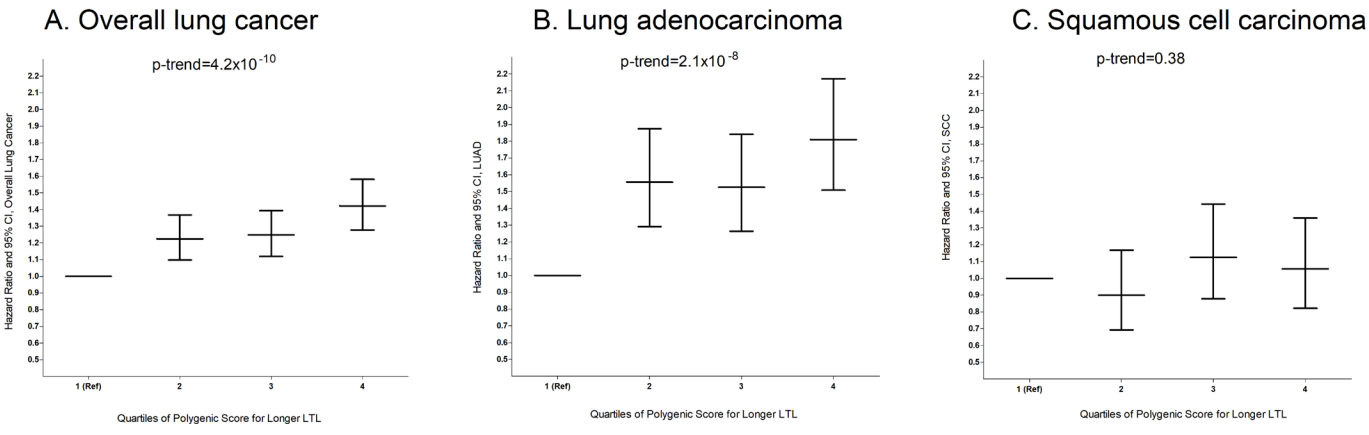

Figure 1 Genetically predicted leucocyte telomere length and risk of lung cancer and its histological subtypes. Using European data from the UK Biobank, we constructed a weighted PGS that predicts longer LTL. This PGS was based on 130 out of 197 genetic variants identified in the UK Biobank that reached genome-wide significance ($p<8.31\times10^{-9}$). These genetic variants were included in the previous MR analyses and were LD pruned ($R^2>0.01$) (Codd et al. *Nat Genet* 2021). The weights were based on the previously reported Z-standardised beta estimates for SNP-LTL associations (Codd et al. *Nat Genet* 2021). We categorised the PGS into quartiles and analysed associations with risk of lung cancer, LUAD, and SCC among Europeans using multivariable Cox regression adjusted for sex, body mass index, age at recruitment, detailed smoking history/intensity, assessment centre, Townsend Deprivation Index, alcohol use, and secondhand smoke exposure. Linear trends were estimated by analysing continuous PGS.

which allows accumulation of genetic alterations that initiate or promote carcinogenesis.

Twitter Mohammad L Rahman @MohammadLRahma2

Acknowledgements We thank Sharon Savage (NCI) for valuable expertise on telomere biology. Additionally, we thank Richard Cawthon for developing the original qPCR assay to measure relative telomere length and for nearly two decades of collaboration. Additionally, we would like to thank Lisa Finkelstein (NCI) for leading efforts to establish the international material transfer agreements.

Contributors JYYW wrote the manuscript, analyzed data, and led the study. BB, AKH, JS, WH, and MLR edited the manuscript and analyzed data. MJM and SMG edited the manuscript. NR and QL edited the manuscript and supervised the study.

Funding This study was supported by intramural funding from the National Cancer Institute (NCI), Division of Cancer Epidemiology and Genetics (DCEG).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by North West - Haydock Research Ethics Committee 3rd Floor - Barlow House 4 Minshull Street Manchester M1 3DZ Telephone: 02071048103. Title of the Database: UK Biobank: a large scale prospective epidemiological resource. Designated Individual: Mrs. Samantha Welsh. REC reference: 21/NW/0157IRAS. project ID: 299116 Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Jason YY Wong http://orcid.org/0009-0007-3690-1168
Mohammad L Rahman http://orcid.org/0000-0001-8322-0495

REFERENCES

1 Codd V, Wang Q, Allara E, *et al*. Polygenic basis and BIOMEDICAL consequences of telomere length variation. *Nat Genet* 2021;53:1425–33.
2 Taub MA, Conomos MP, Keener R, *et al*. Genetic determinants of Telomere length from 109,122 Ancestrally diverse whole-genome sequences in topmed. *Cell Genom* 2022;2:100084.
3 Sudlow C, Gallacher J, Allen N, *et al*. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
4 Codd V, Denniff M, Swinfield C, *et al*. Measurement and initial characterization of Leukocyte Telomere length in 474,074 participants in UK Biobank. *Nat Aging* 2022;2:170–9.
5 Seow WJ, Cawthon RM, Purdue MP, *et al*. Telomere length in white blood cell DNA and lung cancer: a pooled analysis of three prospective cohorts. *Cancer Res* 2014;74:4090–8.
6 Yuan J-M, Beckman KB, Wang R, *et al*. Leukocyte telomere length in relation to risk of lung adenocarcinoma incidence: findings from the Singapore Chinese Health study. Int J Cancer 2018;142:2234–43.
7 Doherty JA, Grieshober L, Houck JR, *et al*. Nested case-control study of telomere length and lung cancer risk among heavy smokers in the β-carotene and retinol efficacy trial. Br J Cancer 2018;118:1513–7.
8 , Haycock PC, Burgess S, *et al*, Telomeres Mendelian Randomization Collaboration. Association between telomere length and risk of cancer and non-neoplastic diseases: a Mendelian randomization study. JAMA Oncol 2017;3:636–51.
9 Kachuri L, Saarela O, Bojesen SE, *et al*. Mendelian randomization and mediation analysis of leukocyte telomere length and risk of lung and head and neck cancers. Int J Epidemiol 2019;48:751–66.
10 Machiela MJ, Hsiung CA, Shu X-O, *et al*. Genetic variants associated with longer Telomere length are associated with increased lung cancer risk among never-smoking women in Asia: a report from the female lung cancer consortium in Asia. *Int J Cancer* 2015;137:311–9.